

A Framework for Generalising the Newton Method and Other Iterative Methods from Euclidean Space to Manifolds

Jonathan H. Manton

*Department of Electrical and Electronic Engineering
The University of Melbourne, Victoria 3010, Australia*

Abstract

The Newton iteration is a popular method for minimising a cost function on Euclidean space. Various generalisations to cost functions defined on manifolds appear in the literature. In each case, the convergence rate of the generalised Newton iteration needed establishing from first principles. The present paper presents a framework for generalising iterative methods from Euclidean space to manifolds that ensures local convergence rates are preserved. It applies to any (memoryless) iterative method computing a coordinate independent property of a function (such as a zero or a local minimum). All possible Newton methods on manifolds are believed to come under this framework. Changes of coordinates, and not any Riemannian structure, are shown to play a natural role in lifting the Newton method to a manifold. The framework also gives new insight into the design of Newton methods in general.

Keywords: Newton iteration, Newton method, convergence rates, optimisation on manifolds, geometric computing

1. Introduction

The Newton iteration function $N_f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ associated with a smooth cost function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is

$$N_f(x) = x - [H_f(x)]^{-1} \nabla f(x), \quad x \in \mathbb{R}^n \quad (1)$$

where $\nabla f(x)$ and $H_f(x)$ are the gradient and Hessian of f , respectively; N_f does not depend on the choice of inner product with respect to which the gradient and Hessian are defined. Starting with an initial guess $x_0 \in \mathbb{R}^n$, the Newton method uses the Newton iteration function to generate the iterates $x_{k+1} = N_f(x_k)$. Under certain conditions [18], this sequence is well-defined and converges to a

Email address: j.manton@ieee.org (Jonathan H. Manton)

critical point of f , meaning $[H_f(x_k)]^{-1}$ exists for all k , and $x = \lim_{k \rightarrow \infty} x_k$ exists and satisfies $\nabla f(x) = 0$.

Let $f: M \rightarrow \mathbb{R}$ now be a smooth cost function defined on an n -dimensional manifold M . Since M locally looks like \mathbb{R}^n , it is natural to ask how the Newton iteration function (1) can be extended to an iteration function $E_f: M \rightarrow M$ such that the iterates $x_{k+1} = E_f(x_k)$ enjoy the same locally quadratic rate of convergence as do the Euclidean Newton iterates.

One approach [11] is to endow the manifold M with a metric and define E_f by a formula analogous to (1) but with ∇f and H_f replaced by the Riemannian gradient $\text{grad } f$ and Hessian \mathcal{H}_f of f , and the straight-line increment $-[H_f(x)]^{-1}\nabla f(x)$ replaced by an increment along a geodesic, namely

$$E_f(p) = \text{Exp}_p \left(-[\mathcal{H}_f(p)]^{-1} \text{grad } f(p) \right) \quad (2)$$

where Exp_p is the Riemannian exponential map centred at p .

The Riemannian Newton method (2) has some disadvantages and other Newton methods on manifolds are possible [5, 15].

What is the most general form of a Newton method on a manifold? Here, a Newton method is defined as any iterative algorithm $p_{k+1} = E_f(p_k)$ that converges locally quadratically to every non-degenerate critical point of every reasonable cost function f , where $E_f(p)$ depends only on the 2-jet of the function f at p ; if f and g agree to second order at p then $E_f(p) = E_g(p)$.

Theorem 11 affords an answer, expressed in terms of parametrisations. A parametrisation of a manifold M is a function $\phi: TM \rightarrow M$ whose restriction $\phi_p: T_pM \rightarrow M$ to the tangent space T_pM of M at any point $p \in M$ provides a (not necessarily one-to-one) correspondence between a neighbourhood of $0_p \in T_pM$ and a neighbourhood of $p \in M$; the former is a subset of a vector space and therefore easier to work with. It suffices for ϕ to be \mathcal{C}^2 -smooth and satisfy $\phi_p(0_p) = p$, but interestingly, there exist valid parametrisations that are not continuous. (Precise definitions are given in the body of the paper.)

Theorem 11 states that for any pair of parametrisations $\phi: TM \rightarrow M$ and $\psi: TM \rightarrow M$, the iteration function $E_f(p) = \psi_p \circ N_{f \circ \phi_p}(0_p)$ is a Newton method on the manifold M , where N is the Euclidean Newton iteration function (1) but on the abstract vector space T_pM rather than \mathbb{R}^n . (Since (1) does not depend on the choice of inner product, there is no need for a Riemannian metric on M .) Justification is given in the body of the paper for believing this to be the most general form possible of a Newton method on a manifold.

Requiring a Newton method to be strictly of the form $p_{k+1} = E_f(p_k)$ places an unnecessary global topological constraint on the parametrisations. Instead, $\phi_p: T_pM \rightarrow M$ could be constructed on demand by “transporting” the old parametrisation $\phi_{p_{k-1}}$ from p_{k-1} to p_k . As transport is generally path dependent, $E_f(p_k)$ may depend on where p_k is relative to p_{k-1} . A uniformity constraint on the family of possible parametrisations allows for the generalisation of Theorem 11 to this situation; see Section 6 for details.

The expression $E_f(p) = \psi_p \circ N_{f \circ \phi_p}(0_p)$ “lifts” the Newton iteration function N from Euclidean space to a manifold. Section 8 explores in generality the lifting of an iteration function from Euclidean space to a manifold.

1.1. Implications, Limitations and Examples

Two broad types of optimisation problems can be distinguished. One is when little is known in advance about the possible cost functions (save perhaps that they are convex, for example) and an algorithm is desired that scales well with increasing dimension. The other is when the family of possible cost functions is known in advance and an algorithm is desired that works well for all members of the family. The latter is the implicit focus of the current paper and relates to real-time optimisation problems in signal processing: at each instance, a new observation y is made; this serves to select a cost function $f(\cdot; y)$; it is required to find quickly an x that maximises $f(x; y)$.

Although generic choices are possible of the pair of parametrisations ϕ and ψ defining a Newton method on a particular manifold, the fact remains that for large-scale problems, the Newton method is generally abandoned in favour of quasi-Newton methods that build up approximations of the Hessian over time, thereby making computational savings by not evaluating the Hessian at each iteration. Quasi-Newton methods have memory and thus are not of the form $x_{k+1} = N_f(x_k)$. An intended sequel will study how to lift algorithms with memory to manifolds.

How can a Newton method be customised for a given family of cost functions? It is propounded that thinking in terms of parametrisations ϕ and ψ offers greater insight into the design of optimisation algorithms. Notwithstanding that identifying a “killer application” for the theory is work in progress, the following example may sway some readers.

Generalising the Rayleigh quotient to higher dimensions yields two well-studied optimisation problems [12]. Recall that the (n, p) -Stiefel manifold is the set of matrices $X \in \mathbb{R}^{n \times p}$ satisfying $X^T X = I$, where superscript T denotes transpose and I is the identity matrix. (The manifold structure is inherited from $\mathbb{R}^{n \times p}$.) Let $A \in \mathbb{R}^{n \times n}$ be symmetric and $N \in \mathbb{R}^{p \times p}$ diagonal, both with distinct positive eigenvalues. A minimising $X \in \mathbb{R}^{n \times p}$ of $f(X) = \text{Tr}(X^T A X N)$ subject to $X^T X = I$ has as its columns the eigenvectors of A corresponding to the p smallest eigenvalues of A . If it was only required to find the subspace spanned by these minor eigenvectors, known as the minor subspace of A , then it suffices to minimise $g(X) = \text{Tr}(X^T A X)$ on the Grassmann manifold. The Grassmann manifold is a quotient space obtained from the Stiefel manifold by declaring two matrices X, Y as equivalent whenever there exists an orthogonal matrix Q such that $Y = XQ$. In other words, each point on the (n, p) -Grassmann manifold represents a particular p -dimensional subspace of \mathbb{R}^n .

The rate of convergence of $X_{k+1} = E_f(X_k) = \psi_{X_k} \circ N_{f \circ \phi_{X_k}}(0_{X_k})$ is dictated by how close to being quadratic $f \circ \phi_X$ is about 0_X whenever X is near a critical point of f . As $f(X) = \text{Tr}(X^T A X N)$ is already quadratic, the parametrisation ϕ_X should be as linear as possible. One possibility is defining $\phi_X(Z)$ as the point on the Stiefel manifold closest (in the Euclidean metric) to the matrix Z ; Section 7.4 proves that parametrisations based on projections are linear to at least second order. The role of ψ is to map $N_{f \circ \phi_X}(0_X)$ back to the manifold with a minimum of fuss. Choosing ψ to be the same as ϕ suffices. (The

option exists of choosing ψ to be an approximation of ϕ that makes $E_f(X) = \psi_X \circ N_{f \circ \phi_X}(0_X)$ overall less computationally demanding to evaluate numerically than $E_f(X) = \phi_X \circ N_{f \circ \phi_X}(0_X)$.) Since the Grassmann manifold is a quotient of the Stiefel manifold, the above argument readily extends to minimising $g(X)$ on a Grassmann manifold; see [15] for the precise calculations.

The above algorithm was trivial to derive yet is a sound starting point upon which clever refinements are possible [2, 4]. The cubic rate of convergence is readily explained in terms of $g \circ \phi_X$ being quadratic to third order at critical points; compare with Example 10. A feature of the derivation is choosing ϕ_X with purpose rather than by trial and error.

How should a theory of optimisation on manifolds be framed? This third italicised objective of the paper is in response to misconceptions including: a connection is required for a Newton method to be definable; only Riemannian Newton methods are “true” Newton methods; and, methods not exploiting the curvature of the manifold must be inferior. These misconceptions come from overplaying the geometry of the manifold itself.

The most relevant geometry is that of the family of cost functions [16]. Knowing the possible cost functions f allows for the customisation of the Newton method by choosing a parametrisation ϕ that makes $f \circ \phi$ approximately quadratic, and such a choice depends not on the manifold M but on the family of cost functions. (Placing a sensible geometry on M might be advantageous — perhaps computational burden can be reduced by exploiting symmetry — but the overall benefit nevertheless will depend on the cost functions.)

It is not pragmatic to insist that only Riemannian Newton methods (2) are true Newton methods. Different methods work better for some cost functions and worse for others; no single method can be superior for every smooth cost function. Any method achieving a locally quadratic rate of convergence is worthy of the title Newton method, provided of course it depends only on the 2-jet of the function; see the definition given earlier.

Under this more general definition, there are Newton methods that cannot be defined in terms of a connection. A connection must vary smoothly whereas no such requirement exists for the parametrisation ϕ . More importantly, thinking of parametrisations instead of connections is more conducive to customising a Newton method for a given family of cost functions. (The Riemannian approach (2) does not offer explicit insight into which metric to use if there are two or more competing metrics, or what to do if there is no convenient choice of metric.)

This paper avoids any need of Riemannian geometry by framing the theory of optimisation on manifolds in terms of robustness of the iteration function to changes of coordinates; see Section 8 for details. This appears to be the most natural point of view.

1.2. Motivation and Relationship with Other Work

Given the extensive background and bibliography made available in the book [1], only a handful of papers are discussed below.

The Riemannian Newton method (2) was introduced in [11] but apparently went unnoticed. The same methodology was rediscovered in the influential

paper [9]. The mindset is that the Newton method is defined by its formula (1), and its extension to a manifold thus necessitates endowing the manifold with a Riemannian metric so the gradient and Hessian can be defined.

Numerically evaluating the Riemannian exponential map in (2) can be costly. It is common to replace the exponential map by an approximation that is cheaper to evaluate numerically. This is formalised in [5], with a precursor in [19]. It corresponds to $E_f(p) = \psi_p \circ N_{f \circ \text{Exp}_p}(0_p)$. A Riemannian metric is still required for computing the Newton increment, with what is termed a *retraction* ψ mapping the result back to the manifold. The retraction ψ must satisfy several conditions, including being smooth. In the present paper, ψ need not be continuous and hence is not even a retraction in the topological sense. The way retractions are commonly used in topology differs in spirit from how ϕ and ψ are being used to lift the Newton method to a manifold, hence the persistence here of calling them parametrisations.

The Riemannian mindset was challenged in [15]. The basic idea is that since the Newton method is a local method, the cost function in a neighbourhood of the current point can be pulled back to a cost function on Euclidean space via a parametrisation, one step of the Newton method carried out in Euclidean space, and the result mapped back to the manifold. No Riemannian metric is necessary. This corresponds to $E_f(p) = \phi_p \circ N_{f \circ \phi_p}(0_p)$. Using projections to define the parametrisations ϕ_p was emphasised. (The resulting algorithms differ significantly from projected Newton methods that take a Newton step in the ambient space then project back to the constraint surface.)

Combining the use of ψ in [5] and the use of ϕ in [15] immediately yields the general form $E_f(p) = \psi_p \circ N_{f \circ \phi_p}(0_p)$ that is the protagonist of the present paper. This form is developed systematically in Sections 4 and 5 in a way that suggests it is the most general form possible of a Newton method.

The use of projections to define parametrisations, advocated in [15], was studied in [3], but for $E_f(p) = \psi_p \circ N_{f \circ \text{Exp}_p}(0_p)$. Convergence proofs were based on calculus techniques requiring more orders of differentiability than necessary; see Section 3.

Another active stream of research is finding lower bounds on the radius of convergence of Riemannian Newton methods [6, 7, 10]. This has not been addressed in the present paper, although in principle, a careful study of the constants in the bounds derived here would provide that information.

The question of the most general form of a Newton method on a manifold appears not to have been addressed before.

2. Basic Notation and Definitions

For a function f between Euclidean spaces, the following definitions are made. The Euclidean norm $\|\cdot\|$ on \mathbb{R}^n is used throughout. The norm of the second-order derivative $D^2f(x)$ is $\|D^2f(x)\| = \sup_{\|\eta\|=1} \|D^2f(x) \cdot (\eta, \eta)\|$. All other norms are operator norms. Gradients ∇f and Hessians H_f are calculated with respect to the Euclidean inner product. The identity operator is denoted

by I (or sometimes by 1 in the one-dimensional case). The notation $B_n(x; \rho)$ and its abbreviation $B(x; \rho)$ denote the open ball centred at $x \in \mathbb{R}^n$ of radius ρ . Its closure is $\overline{B}(x; \rho)$.

An iteration function $N: \mathbb{R}^n \rightarrow \mathbb{R}^n$, which may not be defined on the whole of \mathbb{R}^n , is said to converge locally to $x^* \in \mathbb{R}^n$ with rate $K \in \mathbb{R}$ and constant $\kappa \in \mathbb{R}$ if there exists an open set $U \subset \mathbb{R}^n$ containing x^* such that N is defined on U and

$$\forall x \in U, \quad N(x) \in U \text{ and } \|N(x) - x^*\| \leq \kappa \|x - x^*\|^K. \quad (3)$$

If $K = 1$ then it is further required that $\kappa < 1$, and convergence is called linear. If $K \in (1, 2)$ the convergence is super-linear, and if $K = 2$ the convergence is quadratic.

Although (3) implies $N(x^*) = x^*$, the sequence $x_{k+1} = N(x_k)$ need not converge to x^* for an arbitrary $x_0 \in U$. Nevertheless, define $\bar{\rho} = \kappa^{1/(1-K)}$ if $K > 1$, or $\bar{\rho} = \infty$ if $K = 1$ and $\kappa < 1$. Then $B(x^*; \rho) \cap U$ is mapped into itself by N whenever $\rho \leq \bar{\rho}$. Moreover, $x_0 \in B(x^*; \bar{\rho}) \cap U$ implies $x_k \rightarrow x^*$.

The focus of this paper is on convergence rates K greater than one.

3. Local Convergence of the Newton Iteration on Euclidean Space

Convergence proofs for the Newton method include the Newton-Kantorovich theorem (applicable for the Newton method on Banach spaces) and the Newton-Mysovskikh theorem; see [14, 17] and the bibliographic note [17, p. 428]. These theorems give sufficient but not necessary conditions, concentrating instead on explicitly finding a region within which the Newton method is guaranteed to converge. The affine invariance of the Newton method is exploited in [8] to sharpen these classical results.

In pursuit of the most general Newton method on a manifold, it is informative to derive a necessary and sufficient condition for the standard Newton method to converge to a non-degenerate critical point.

Theorem 1. *Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^2 -smooth. Let $x^* \in \mathbb{R}^n$ be a non-degenerate critical point, that is, $\nabla f(x^*) = 0$ and $H_f(x^*)$ is invertible. A necessary and sufficient condition for N_f in (1) to be locally quadratically convergent to x^* is for there to exist $\eta, \delta > 0$ such that $x \in B(x^*; \delta)$ implies*

$$\| [H_f(x) - H_f(x^*)](x - x^*) \| \leq \eta \|x - x^*\|^2. \quad (4)$$

Proof. Define the second-order Taylor series remainder term

$$R(x) = f(x) - f(x^*) - \frac{1}{2}(x - x^*)^T H_f(x^*)(x - x^*). \quad (5)$$

Since f is \mathcal{C}^2 , so is R . Moreover,

$$\nabla f(x) = H_f(x^*)(x - x^*) + \nabla R(x), \quad (6)$$

$$H_f(x) = H_f(x^*) + H_R(x). \quad (7)$$

Substitution into (1) shows

$$N_f(x) - x^* = x - x^* - [H_f(x)]^{-1} [H_f(x^*)(x - x^*) + \nabla R(x)] \quad (8)$$

$$= [H_f(x)]^{-1} [H_R(x)(x - x^*) - \nabla R(x)]. \quad (9)$$

Since H_R is continuous, for any $\epsilon > 0$ there exists a $\rho > 0$ such that $x \in B(x^*; \rho)$ implies: $H_f(x)$ is invertible; $\|[H_f(x)]^{-1}\| \leq \|[H_f(x^*)]^{-1}\| + \epsilon$; and $\|H_f(x)\| \leq \|H_f(x^*)\| + \epsilon$.

To prove sufficiency, first observe

$$\|N_f(x) - x^*\| \leq \|[H_f(x)]^{-1}\| (\|H_R(x)(x - x^*)\| + \|\nabla R(x)\|). \quad (10)$$

Choose δ, η as in the theorem. If $x \in B(x^*; \delta)$ then

$$\|\nabla R(x)\| \leq \int_0^1 \frac{1}{t} \|H_R(x^* + t(x - x^*))t(x - x^*)\| dt \quad (11)$$

$$\leq \int_0^1 t \eta \|x - x^*\|^2 dt \quad (12)$$

$$\leq \frac{1}{2} \eta \|x - x^*\|^2. \quad (13)$$

Choosing ϵ, ρ as above, if $x \in B(x^*; \min\{\delta, \rho\})$ then $N_f(x)$ is well-defined and

$$\|N_f(x) - x^*\| \leq \frac{3}{2} \eta (\|[H_f(x^*)]^{-1}\| + \epsilon) \|x - x^*\|^2, \quad (14)$$

proving local quadratic convergence.

To prove necessity, first note from (9) that

$$\|N_f(x) - x^*\| \geq \|H_f(x)\|^{-1} \|H_R(x)(x - x^*) - \nabla R(x)\|. \quad (15)$$

Thus, choosing ϵ, ρ as above, if $x \in B(x^*; \rho)$ then

$$\|H_R(x)(x - x^*) - \nabla R(x)\| \leq (\|H_f(x^*)\| + \epsilon) \|N_f(x) - x^*\|. \quad (16)$$

By hypothesis, N_f converges locally quadratically to x^* , hence by shrinking ρ if necessary, there exists a $\kappa > 0$ such that $x \in B(x^*; \rho)$ implies

$$\|H_R(x)(x - x^*) - \nabla R(x)\| \leq \kappa \|x - x^*\|^2. \quad (17)$$

Define the closed ball $C = \overline{B}(x^*; \rho/2)$ and the function $\phi(x) = \|H_R(x)(x - x^*)\| \|x - x^*\|^{-1}$. Setting $\phi(x^*) = 0$ ensures ϕ is well-defined and continuous on C . Assume to the contrary, for all $\eta > 0$, the scalar $h = \max_{x \in C} \{\phi(x) - \eta \|x - x^*\|\}$

satisfies $h > 0$. For any $x \in C$,

$$\|\nabla R(x)\| \leq \int_0^1 \|H_R(x^* + t(x - x^*))(x - x^*)\| dt \quad (18)$$

$$= \|x - x^*\| \int_0^1 \phi(x^* + t(x - x^*)) dt \quad (19)$$

$$\leq \|x - x^*\| \int_0^1 h + t\eta \|x - x^*\| dt \quad (20)$$

$$= h\|x - x^*\| + \frac{1}{2}\eta\|x - x^*\|^2. \quad (21)$$

Let $z \in C$ be such that $\phi(z) - \eta\|z - x^*\| = h$. Since $z \neq x^*$ and

$$\|H_R(z)(z - x^*) - \nabla R(z)\| \geq \phi(z)\|z - x^*\| - h\|z - x^*\| - \frac{1}{2}\eta\|z - x^*\|^2 \quad (22)$$

$$= \frac{1}{2}\eta\|z - x^*\|^2, \quad (23)$$

choosing $\eta > 2\kappa$ makes (23) contradict (17), proving the theorem. \square

Corollary 2. *Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be C^3 -smooth and $x^* \in \mathbb{R}^n$ a non-degenerate critical point. Then N_f in (1) converges locally quadratically to x^* .*

Proof. If f is C^3 then $H_f(x) - H_f(x^*)$ is C^1 , hence (4) holds. \square

Corollary 3. *Let f and x^* satisfy the conditions in Theorem 1, including (4). The perturbed iteration function $E_f(x) = x - [H_f(x) + G(x)]^{-1}\nabla f(x)$ converges locally quadratically to x^* if there exists a $\gamma \in \mathbb{R}$ such that the operator norm of the matrix $G(x)$ satisfies $\|G(x)\| \leq \gamma\|x - x^*\|$ in a neighbourhood of x^* .*

Proof. Observe

$$E_f(x) - x^* = [H_f(x) + G(x)]^{-1} \{H_f(x)(x - x^*) - \nabla f(x) + G(x)(x - x^*)\}. \quad (24)$$

Therefore,

$$\|E_f(x) - x^*\| \leq \|[H_f(x) + G(x)]^{-1}\| \{ \|\nabla f(x) - H_f(x^*)(x - x^*)\| + \|[H_f(x) - H_f(x^*)](x - x^*)\| + \|G(x)(x - x^*)\| \}. \quad (25)$$

In a sufficiently small neighbourhood of x^* , $\|[H_f(x) + G(x)]^{-1}\|$ is bounded above by a constant and the three other terms are bounded by a constant times $\|x - x^*\|^2$; refer to (13) and the hypotheses on $H_f(x)$ and $G(x)$. \square

Despite calculus offering a simpler and more elegant alternate, convergence proofs are based here on hard analysis because calculus requires a higher order of smoothness than necessary, as now demonstrated. (See also the opening paragraph of Appendix C.2.) Recall the basic principle.

Lemma 4. Let $N: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be \mathcal{C}^K -smooth for some integer $K \geq 2$. If $D^k N(x^*) = 0$ for $k = 1, \dots, K-1$ then N converges locally to x^* with rate K .

Applying Lemma 4 to (1) shows that f being \mathcal{C}^4 -smooth is sufficient for N_f to converge locally quadratically to a nondegenerate critical point. If f were only \mathcal{C}^3 then N_f would only be \mathcal{C}^1 and Lemma 4 could not be applied. The actual condition (4) falls strictly between \mathcal{C}^2 -smoothness and \mathcal{C}^3 -smoothness.

Example 5. Define $f(x) = x^2 + |x|^{5/2}$. The origin is a non-degenerate critical point. The Newton iteration function is $N_f(x) = \frac{5x|x|^{1/2}}{8+15|x|^{1/2}}$ and has super-linear but not quadratic convergence, despite f being \mathcal{C}^2 -smooth.

Remark 6. The quadratic convergence rate of the Newton method is coordinate independent, in the following sense. Assume $f: \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the conditions in Theorem 1 about the point x^* . If $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a \mathcal{C}^2 -diffeomorphism then $\phi^{-1}(x^*)$ is a non-degenerate critical point of $f \circ \phi$, and by Proposition 37, condition (4) holds for $f \circ \phi$ about the point $\phi^{-1}(x^*)$. Thus, if N_f converges locally quadratically to x^* then $N_{f \circ \phi}$ converges locally quadratically to $\phi^{-1}(x^*)$.

4. The Coordinate Adapted Newton Iteration

The most general form of a Newton method in Euclidean space is explored.

4.1. Coordinate Adaptation

Applying a change of coordinates $\phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ to (1) yields the new iteration function $E_f(x) = \phi \circ N_{f \circ \phi} \circ \phi^{-1}(x)$. Expedient choices of ϕ can increase the domain of attraction, decrease the computational complexity per iteration and improve the convergence rate. As an extreme example, if ϕ is such that $f \circ \phi$ is quadratic then E_f converges in a single iteration. Although Morse's Lemma guarantees the existence of such a ϕ locally, finding it is generally not practical. This motivates using a different change of coordinates at each iteration, namely $E_f(x) = \phi_x \circ N_{f \circ \phi_x} \circ \phi_x^{-1}(x)$. When ϕ_x varies with x , the convergence properties of E_f need not follow from the convergence properties of N_f . Significantly then, it is established that under mild conditions, E_f converges locally quadratically to non-degenerate critical points of f .

Coordinate adaptation is defined in terms of a function $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, alternatively written $\phi_x(y) = \phi(x, y)$, satisfying the condition that, $\forall x^* \in \mathbb{R}^n$, $\exists \alpha, \beta, \rho \in \mathbb{R}$, $\rho > 0$, $\forall x, y \in B(x^*; \rho)$, the following hold:

P1 $\|D^2\phi_x(x)\| \leq \alpha$;

P2 $\|\phi_x(y) - y\| \leq \beta\|y - x\|^2$.

Implicit in P1 is the requirement that $D^2\phi_x(x)$ exists, which in turn requires the existence of $D\phi_x(y)$ for y sufficiently close to x .

Given such a ϕ , the coordinate adapted Newton iteration function is

$$E_f(x) = \phi_x \circ N_{f \circ \phi_x}(x) \tag{26}$$

where N_f is the Newton iteration function (1). This agrees with the earlier expression for E_f because P2 implies $\phi_x(x) = x$.

Theorem 7. *Let f and x^* satisfy the conditions in Theorem 1, including (4). Let ϕ satisfy P1 and P2, defined above. Then the coordinate adapted Newton iteration function E_f , defined in (26), converges locally quadratically to x^* .*

Proof. P2 implies $\phi_x(x) = x$ and $D\phi_x(x) = I$. Hence

$$D(f \circ \phi_x)(x) = Df(x), \quad (27)$$

$$D^2(f \circ \phi_x)(x) = D^2f(x) + Df(x)D^2\phi_x(x). \quad (28)$$

Let $G(x)$ be the matrix representation of $Df(x)D^2\phi_x(x)$. Then $G(x)$ is symmetric and satisfies $\xi^T G(x)\xi = Df(x)D^2\phi_x(x) \cdot (\xi, \xi)$ for any $\xi \in \mathbb{R}^n$. If P1 holds, it can be shown that $\|G(x)\| \leq \alpha\|Df(x)\|$. Thus, in a neighbourhood of x^* , there exist constants $r, \gamma > 0$ such that $x \in B(x^*; r)$ implies $\|G(x)\| \leq \gamma\|x - x^*\|$. Since $H_{f \circ \phi_x}(x) = H_f(x) + G(x)$, it follows from Corollary 3 that, for a possibly smaller $r > 0$, there exists a $\kappa > 0$ such that $\|N_{f \circ \phi_x}(x) - x^*\| \leq \kappa\|x - x^*\|^2$ whenever $x \in B(x^*; r)$. To be able to apply P2, shrink r if necessary to ensure $0 < r \leq \rho$. Then

$$\|E_f(x) - x^*\| \leq \|\phi_x(N_{f \circ \phi_x}(x)) - N_{f \circ \phi_x}(x)\| + \|N_{f \circ \phi_x}(x) - x^*\| \quad (29)$$

$$\leq \beta\|N_{f \circ \phi_x}(x) - x\|^2 + \kappa\|x - x^*\|^2 \quad (30)$$

$$\leq \beta(\|N_{f \circ \phi_x}(x) - x^*\| + \|x - x^*\|)^2 + \kappa\|x - x^*\|^2 \quad (31)$$

$$\leq (\beta(\kappa r + 1)^2 + \kappa)\|x - x^*\|^2 \quad (32)$$

whenever $x \in B(x^*; r)$, proving the theorem. \square

As now explained, P1 and P2 are not only mild, it is conjectured they cannot be weakened. For (26) to be defined, the Hessian of $f \circ \phi_x$ at x must exist, necessitating the existence of $D^2\phi_x(x)$ implicit in P1. The local bound in P1 ensures $N_{f \circ \phi_x}(x)$ converges locally quadratically. A side-effect of P2 is that $\phi_x(x) = x$ and $D\phi_x(x) = I$; this loses no generality because the Newton iteration function is invariant to affine changes of coordinates. The main purpose of P2 is to prevent the residual term $R_x(y) = \phi_x(y) - y - \frac{1}{2}D^2\phi_x(x) \cdot (y - x, y - x)$ from being unbounded locally, ensuring that if $E(x)$ is an arbitrary iteration function converging locally quadratically to x^* then $\phi_x \circ E(x)$ continues to converge locally quadratically to x^* . The situation in which, for a sufficiently large class of cost functions f , $N_{f \circ \phi_x}(x)$ fails to have local quadratic convergence yet E_f has local quadratic convergence is conjectured to be impossible. The claim that P1 and P2 are mild comes from the fact deducible from Appendix B that any \mathcal{C}^2 -smooth ϕ with $\phi_x(x) = x$ and $D\phi_x(x) = I$, satisfies P1 and P2. Furthermore, neither $\phi(x, y)$ or $D\phi_x(y)$ need be continuous except on the diagonal $y = x$, and ϕ_x need not be locally \mathcal{C}^1 -smooth.

The following example shows that if arbitrary changes of coordinates are allowed then the coordinate adapted Newton method may not even be defined, much less converge at a quadratic rate.

Example 8. Let β be an arbitrary scalar. Consider the coordinate adapted Newton iteration function applied to $f(x) = x^2$ using $\phi_x(y) = y + \frac{\beta}{x}(y-x)^2$ when $x \neq 0$ and $\phi_x(y) = y$ when $x = 0$. If $x \neq 0$ then $N_{f \circ \phi_x}(x) = \frac{2\beta}{1+2\beta}x$ which is not defined if $\beta = -\frac{1}{2}$. If $\beta \neq -\frac{1}{2}$ then $E_f(x) = \frac{(3+4\beta)\beta}{(1+2\beta)^2}x$, which in general exhibits at best linear convergence.

4.2. The Generalised Coordinate Adapted Newton Iteration

The proof of Theorem 7 can be modified trivially to prove the following result. A new function ψ analogous to ϕ is introduced and property P2 in Section 4.1 is replaced by

$$\mathbf{P2'} \quad \|\psi_x(y) - y\| \leq \beta \|y - x\|^2.$$

Theorem 9. Let f and x^* satisfy the conditions in Theorem 1, including (4). Let $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfy P1 in Section 4.1. Assume further that $\phi_x(x) = x$ and $D\phi_x(x) = I$. Let $\psi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfy P2' above; the qualifiers for x and y in P2' are the same as for P2. Then the generalised coordinate adapted Newton iteration function

$$E_f(x) = \psi_x \circ N_{f \circ \phi_x}(x) \tag{33}$$

converges locally quadratically to x^* .

Being able to change both ϕ and ψ allows greater control over the computational complexity, the domain of attraction and the rate of convergence of the iteration function (33), as now discussed.

4.3. Discussion

The choice of coordinate changes ϕ_x and ψ_x in (33) determines which class of cost functions the generalised coordinate adapted Newton method will perform well for. The challenge then is to determine suitable coordinate changes to use for the class of cost functions at hand. For inherently difficult optimisation problems this will not be easy by definition. Nevertheless, thinking in terms of coordinate adaptation leads to the following new strategy.

The closer the cost function is to being quadratic, the faster the convergence rate of (1). Ideally then, ϕ_x in (26) makes $f \circ \phi_x$ approximately quadratic for every cost function f in the given family. For improving local convergence, it suffices to restrict attention to cost functions with a critical point near x because, by definition of local, it can be assumed a critical point is nearby x , and this limits the possibilities of which cost function from the family has been selected to be minimised. For the special case when f has the form $f(x) = \psi(x-z)$ for some unknown scalar z , where ψ has a minimum at the origin, it suffices for $f \circ \phi_x$ to be approximately quadratic when $z = x$.

Example 10. Consider the family of cost functions $f(x; z) = (x-z)^2 + 2(x-z)^3$. The coordinate adapted Newton iteration function using the coordinate systems $\phi_x(y) = y - (y-x)^2$ is $E_f(x; z) = z - 8(x-z)^3 + \dots$. This converges cubically to the critical point $x^* = z$ for any cost function in the family. Here, $\phi_x(y)$ was chosen so $f \circ \phi_z(x) = (x-z)^2 - 5(x-z)^4 + \dots$ has no cubic term.

If the domain of attraction is of primary concern then similar intuition suggests choosing ϕ_x such that, for any f belonging to the given class of cost functions, $f \circ \phi_x$ has a relatively large domain of attraction, especially if x is at all close to the minimum of f .

The extra freedom afforded by ψ_x in (33) can be used to reduce the computational complexity per iteration without compromising the rate of convergence; in some cases, an expedient choice of ψ_x leads to cancellations, so $\psi_x \circ N_{f \circ \phi_x}$ becomes less computationally intensive to evaluate than $N_{f \circ \phi_x}$ on its own.

The coordinate adapted Newton method is different from variable metric methods. Variable metric methods explicitly or implicitly perform a change of coordinates and then take a steepest-descent (not Newton) step in the new coordinate system. They do not evaluate the Hessian of the cost function but instead build up an approximation B_k to the Hessian from current and past gradient information. They are of the form $x_{k+1} = x_k - [B_k]^{-1} \nabla f(x_k)$; see [18]. The generalised coordinate adapted Newton method (33) with $\psi_x(y) = y$ can be written as $E_f(x) = x - [H_f(x) + G(x)]^{-1} \nabla f(x)$; see the proof of Theorem 7. This differs from a variable metric method in several ways; E_f makes use of the Hessian of f but B_k does not; E_f has no “memory” but B_k is built up over time; variable metric methods generally only achieve super-linear convergence whereas E_f has quadratic convergence. The philosophy is also different; variable metric methods wish for B_k to be as close as possible to the true Hessian, whereas the generalised coordinate adapted Newton method intentionally uses a perturbed version of the true Hessian to improve the performance of the algorithm.

5. Generalised Newton Methods on Manifolds

Throughout this section, $f: M \rightarrow \mathbb{R}$ will be a \mathcal{C}^2 -smooth cost function defined on an n -dimensional \mathcal{C}^2 -differentiable manifold M . Recall from Appendix A that if $E: M \rightarrow M$ is an iteration function with local quadratic convergence to a point $p^* \in M$ then $\varphi \circ E \circ \varphi^{-1}$ converges locally quadratically to $\varphi(p^*)$ for any chart (U, φ) with $p^* \in U$. Fix (U, φ) . For cost functions f with a critical point in U , the coordinate adapted Newton method of Section 4 can be extended to manifolds by seeking an E_f such that $\varphi \circ E_f \circ \varphi^{-1}$ is a coordinate adapted Newton iteration function for the equivalent cost function $f \circ \varphi^{-1}$. For functions with critical points outside U , in principle a different coordinate chart needs to be taken, but as shown presently, it is straightforward to guess an appropriate form for E_f globally.

Solving $\varphi \circ E_f \circ \varphi^{-1}(x) = \phi_x \circ N_{(f \circ \varphi^{-1}) \circ \phi_x} \circ \phi_x^{-1}(x)$ yields

$$E_f(p) = (\varphi^{-1} \circ \phi_{\varphi(p)}) \circ N_{f \circ (\varphi^{-1} \circ \phi_{\varphi(p)})} \circ (\phi_{\varphi(p)}^{-1} \circ \varphi)(p) \quad (34)$$

$$= \psi_p \circ N_{f \circ \psi_p} \circ \psi_p^{-1}(p), \quad \psi_p(z) = \varphi^{-1} \circ \phi_{\varphi(p)}(z). \quad (35)$$

The affine invariance of the Newton method allows this to be rewritten as $E_f(p) = \tilde{\psi}_p \circ N_{f \circ \tilde{\psi}_p} \circ \tilde{\psi}_p^{-1}(p) = \tilde{\psi}_p \circ N_{f \circ \tilde{\psi}_p}(0)$ where $\tilde{\psi}_p(z) = \psi_p(z + \varphi(p))$. Although this defines E_f only locally, an obvious extension is $E_f(p) = \psi_p \circ$

$N_{f \circ \psi_p}(0)$ where, for each $p \in M$, $\psi_p: \mathbb{R}^n \rightarrow M$ is a parametrisation of a neighbourhood on the n -dimensional manifold M centred at p , that is, $\psi_p(0) = p$. This extension is justified by the proof of Theorem 11 in which it is shown that $\varphi \circ E_f \circ \varphi^{-1}$ does indeed take the form of a coordinate adapted Newton method for any chart φ on M .

Although tempting to generalise ϕ and ψ in Section 4 to maps from $M \times \mathbb{R}^n$ to M , the global geometry of M can prevent any such map from being smooth. The tangent bundle TM , being equivalent to $M \times \mathbb{R}^n$ locally, offers an alternative. As TM twists in the “right” way, smooth parametrisations from TM to M can be anticipated to exist; this was appreciated by Shub [19, 5]. (While smoothness is not essential, in practice it may be convenient to work with smooth parametrisations.)

The functions $\phi: TM \rightarrow M$ and $\psi: TM \rightarrow M$ will be required to satisfy conditions C1–C2 below, which generalise P1 and P2’ in Sections 4.1 and 4.2. Local coordinates are needed. Let $\pi: TM \rightarrow M$ be the projection taking a tangent vector $v_p \in T_pM$ to its base point p . A \mathcal{C}^2 -chart (U, φ) induces the \mathcal{C}^1 -chart $\tau_\varphi: \pi^{-1}(U) \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ on TM , sending v_p to $(\varphi(p), A_p(v_p))$ where $A_p: T_pM \rightarrow \mathbb{R}^n$ is the linear isomorphism taking $v_p = \sum_{j=1}^n \eta^j \frac{\partial}{\partial \varphi_j} \Big|_p$ to $\eta = (\eta^1, \dots, \eta^n)$. The local coordinate representation of ϕ is $\widehat{\phi} = \varphi \circ \phi \circ \tau_\varphi^{-1}$.

Conditions C1–C2 are satisfied if, $\forall p \in M, \exists \mathcal{C}^2$ -chart (U, φ) with $\varphi(p) = 0, \exists \rho > 0$:

C1 $\widehat{\phi} = \varphi \circ \phi \circ \tau_\varphi^{-1}$ satisfies H1 $_\rho$ and H2 $_\rho$, defined below;

C2 $\widehat{\psi} = \varphi \circ \psi \circ \tau_\varphi^{-1}$ satisfies H3 $_\rho$, defined below.

Consider a function $h: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ which need not be defined everywhere; its domain of definition will be clarified shortly. It satisfies **H1** $_\rho$ if $x \in B(0; \rho)$ implies $h_x(0) = x$ and $Dh_x(0) = I$. It satisfies **H2** $_\rho$ if there exists a constant $\alpha \in \mathbb{R}$ such that $\|D^2h_x(0)\| \leq \alpha$ for every $x \in B(0; \rho)$. It satisfies **H3** $_\rho$ if there exists a constant $\beta \in \mathbb{R}$ such that $x, y \in B(0; \rho)$ implies $\|h_x(y) - x - y\| \leq \beta \|y\|^2$. If the subscript ρ is omitted, the existence of an appropriate $\rho > 0$ is implied.

For the derivatives to exist, if h satisfies H1 $_\rho$ or H2 $_\rho$ then its domain of definition must include a set of the form $\{(x, y) \mid x \in B(0; \rho), y \in B(0; \delta_x), \delta_x > 0\}$ where δ_x is a function of x . Such a set need not contain a neighbourhood of the origin. For H3 $_\rho$ though, it is required that $B(0; \rho) \times B(0; \rho)$ lies in the domain of h . See Appendix B for further properties.

A generalised Newton iteration function is any $E_f: M \rightarrow M$ of the form

$$E_f(p) = \psi_p \circ N_{f \circ \phi_p}(0_p) \quad (36)$$

where $\phi_p: T_pM \rightarrow M$ and $\psi_p: T_pM \rightarrow M$ are the restrictions of ϕ and ψ to the tangent space T_pM at the point p on M . In (36), N represents the Newton iteration (1) but on the abstract vector space T_pM rather than \mathbb{R}^n .

The local coordinate representation $\widehat{\phi} = \varphi \circ \phi \circ \tau_\varphi^{-1}$ of ϕ can be written as $\widehat{\phi}(x, y) = \varphi \circ \phi_{\varphi^{-1}(x)} \circ A_{\varphi^{-1}(x)}^{-1}(y)$, alternatively denoted $\widehat{\phi}_x(y)$. The local

coordinate representation of ϕ_p is $\varphi \circ \phi_p \circ A_p^{-1}$, which can be written in terms of $\widehat{\phi}$, namely, $\widehat{\phi}_{\varphi(p)}$. Analogously for ψ_p .

Theorem 11. *Let $f: M \rightarrow \mathbb{R}$ be a \mathcal{C}^2 -smooth cost function on a \mathcal{C}^2 -smooth manifold M . Let $p^* \in M$ be a non-degenerate critical point, that is, $Df(p^*) = 0$ and if $D^2f(p^*) \cdot (\xi, \xi) = 0$ then $\xi = 0$. Assume there is a chart (U, φ) on M , with $\varphi(p^*) = 0$, and an $\eta > 0$ such that $\widehat{f} = f \circ \varphi^{-1}$ satisfies, for $x \in \varphi(U)$,*

$$\|[H_{\widehat{f}}(x) - H_{\widehat{f}}(0)]x\| \leq \eta \|x\|^2. \quad (37)$$

Let $\phi, \psi: TM \rightarrow M$ satisfy C1–C2, defined above. Then the generalised Newton iteration function (36) converges locally quadratically to p^ .*

Proof. Let (U, φ) be as in the theorem. Proposition 38 implies there exists a $\rho > 0$ such that $\widehat{\phi} = \varphi \circ \phi \circ \tau_\varphi^{-1}$ satisfies H1 $_\rho$ and H2 $_\rho$, and $\widehat{\psi} = \varphi \circ \psi \circ \tau_\varphi^{-1}$ satisfies H3 $_\rho$. (By Lemma 31, $\widehat{\psi}$ will also satisfy H1 $_\rho$.) Let $h_x(y) = y - x$, $\tilde{\phi}_x(y) = \widehat{\phi}(x, y - x) = \widehat{\phi}_x \circ h_x(y)$ and $\tilde{\psi}_x(y) = \widehat{\psi}(x, y - x) = \widehat{\psi}_x \circ h_x(y)$. The invariance of the Newton iteration function to the affine coordinate change $h_x^{-1} \circ A_{\varphi^{-1}(x)}$ can be used to show

$$\varphi \circ E_f \circ \varphi^{-1}(x) = \tilde{\psi}_x \circ N_{\widehat{f} \circ \tilde{\phi}_x}(x). \quad (38)$$

The functions $\tilde{\phi}$, $\tilde{\psi}$ and \widehat{f} satisfy the necessary conditions locally about the point $x^* = 0$ for the proof of Theorem 9 to go through. \square

If f is \mathcal{C}^3 -smooth then (37) holds. See also Appendix C.

5.1. A Global Topological Constraint

The generalised Newton method (36) is defined by the choice of parametrisations ϕ and ψ , and Lemma 39 shows it suffices to choose ϕ and ψ to be \mathcal{C}^2 -smooth to ensure local quadratic convergence to non-degenerate critical points. This may appear elegant and straightforward. In practice though, especially if M is a quotient space, directly writing down a smooth parametrisation ϕ may not be the most desirable approach. While numerous possible choices may come to mind for each $\phi_p: T_pM \rightarrow M$, difficulties arise if no *canonical* choice is evident for each p that would make ϕ smooth.

This difficulty is a consequence of a deeper fact: a global topological constraint is unwittingly imposed by insisting that a Newton method be strictly of the form $p_{k+1} = E_f(p_k)$. To see this, consider assigning parametrisations ϕ_p point-by-point on a sphere by starting at a particular point and spreading out in all directions. The non-flatness of the sphere causes these initially divergent directions to begin to converge, with some points ultimately reached from multiple directions. Unless special care is taken, the parametrisations will not match up at such points.

This has not been seen before as a problem because implicit or explicit use typically has been made of a Riemannian metric to guide the construction of

parametrisations. Furthermore, the affine invariance of the Newton method plays a critical role as it means parametrisations constructed locally only have to agree globally with each other up to affine transformations, which is easier (indeed, possible) to achieve. When thinking of a manifold as an object embedded in Euclidean space, it is visually clear how the affine tangent plane can be moved around the manifold, and although two different paths from p to q may move the affine tangent plane differently, the only difference will be a rotation. Strategies such as projection from the affine tangent plane onto the manifold [15], or the use of the Riemannian exponential map, can then be used to generate a smooth parametrisation $\phi: TM \rightarrow M$ with which the Newton method can be lifted from Euclidean space to the manifold M .

The sphere S^2 highlights the role of affine invariance. Although smooth parametrisations $\phi: TS^2 \rightarrow S^2$ are readily constructed, the hairy-ball theorem states there is no global section of the frame bundle on S^2 , meaning there is no way to identify each $T_p S^2$ with \mathbb{R}^2 in a smooth way. If not for the affine invariance of the Newton method, it would be impossible to construct smoothly varying parametrisations with which to lift the Newton method to S^2 . For more general iteration functions, this highlights the importance of conditions C1–C2 not requiring continuity; see Appendix B.

6. Whimsical and Path-Dependent Newton Methods

The global topological constraint in Section 5.1 makes it desirable to allow Newton methods on manifolds to construct their parametrisations dynamically as the sequence of iterates unfolds. Such methods are called path-dependent Newton methods because the parametrisation ϕ_{p_k} used at the k th step may depend on the path p_0, p_1, \dots, p_{k-1} leading up to p_k . One of the many possibilities this opens up is using (non-metric) affine connections and parallel transport to construct parametrisations. See Section 7 for other possibilities.

The theory in Section 5 extends to encompass path-dependent Newton methods because there is no inherent requirement for the ϕ_p to vary smoothly, or even continuously, in p . The essence of H1–H3 in Section 5 is that the bounds α and β hold *uniformly* on sufficiently small neighbourhoods. Therefore, a generalised Newton method can be constructed by specifying a ϕ_p for each p with little regard for how the ϕ_p fit together to form ϕ . In other words, a generalised Newton method at each step is free to choose from many different parametrisations ϕ_p without affecting its performance.

It is expedient to study path-dependent Newton methods in terms of whimsical Newton methods. Let $\Sigma = \{\Sigma_p \mid p \in M\}$ be a collection of sets Σ_p of pairs (ϕ_p, ψ_p) of parametrisations $\phi_p, \psi_p: T_p M \rightarrow M$. A whimsical Newton method with respect to Σ is the general term given to any iterative scheme $p_{k+1} = E_f^k(p_k)$ where $E_f^k(p)$ is a generalised Newton iteration function (36) using a pair of parametrisations (ϕ_p, ψ_p) belonging to Σ_p . Indexing E_f by k means that even if $p_{k+j} = p_k$ for some $j > 0$, a different parametrisation pair can be chosen from Σ_{p_k} at the k th and $(k+j)$ th steps.

Imposing the following uniformity constraints on the elements of Σ ensures any whimsical Newton method with respect to Σ converges locally quadratically to non-degenerate critical points. Recall the definition of A_p in Section 5.

Conditions E0–E2 are satisfied if Σ_p is non-empty for all $p \in M$, and furthermore, $\forall p^* \in M, \exists \mathcal{C}^2$ -chart (U, φ) with $\varphi(p^*) = 0, \exists \rho > 0, \exists \alpha, \beta \in \mathbb{R}, \forall p \in U, \forall (\phi_p, \psi_p) \in \Sigma_p$:

E0 $\phi_p(0_p) = \psi_p(0_p) = p$ and $D\phi_p(0_p) = D\psi_p(0_p) = I$;

E1 $\widehat{\phi}_p = \varphi \circ \phi_p \circ A_p^{-1}$ satisfies $\|D^2\widehat{\phi}_p(0)\| \leq \alpha$;

E2 $\widehat{\psi}_p = \varphi \circ \psi_p \circ A_p^{-1}$ satisfies $\|\widehat{\psi}_p(y) - \varphi(p) - y\| \leq \beta\|y\|^2$ whenever $\|y\| < \rho$.

Theorem 12. *Let $f: M \rightarrow \mathbb{R}$ be a \mathcal{C}^2 -smooth cost function on a \mathcal{C}^2 -smooth manifold M satisfying (37) at a non-degenerate critical point p^* . If Σ satisfies E0–E2 then any whimsical Newton method $p_{k+1} = E_f^k(p_k)$ with respect to Σ converges locally quadratically to p^* . Uniform bounds exist for the rate of convergence (that is, the constants κ and K in (3) and Definition 28 with respect to a given local coordinate chart) that are independent of how $E_f^k(p)$ selects which pair $(\phi_p, \psi_p) \in \Sigma_p$ to use at the k th step.*

Theorem 12 can be proved by observing in the proof of Theorem 7 that the rate of convergence is determined purely in terms of bounds on the second-order behaviour of the parametrisations. Provided the bounds α and β remain valid, the pairs of parametrisations used become irrelevant. Similarly, it follows from the proofs of Proposition 38 and Lemma 36 that if conditions E0–E2 are satisfied with respect to one chart, they are satisfied with respect to any other.

A path-dependent Newton method differs from a whimsical Newton method in that the rule for choosing the parametrisation pair to use at each step may depend on previous iterates.

Corollary 13. *Let $f: M \rightarrow \mathbb{R}$ be a \mathcal{C}^2 -smooth cost function on a \mathcal{C}^2 -smooth manifold M satisfying (37) at a non-degenerate critical point p^* . Let $p_{k+1} = E_f^{(k, p_0, \dots, p_{k-1})}(p_k)$ be a path-dependent Newton iterate with respect to an indexed family $\Sigma = \{\Sigma_p \mid p \in M\}$ of sets Σ_p of parametrisation pairs: $p_{k+1} = \psi_{(k)} \circ N_{f \circ \phi_{(k)}}(0_{p_k})$ where the rule for choosing $(\psi_{(k)}, \phi_{(k)}) \in \Sigma_{p_k}$ may depend on past iterates p_0, \dots, p_{k-1} as well as on k and p_k . If Σ satisfies E0–E2 then this path-dependent Newton iterate converges locally quadratically to p^* .*

Proof. Assume to the contrary the existence of f, p^* and a sequence of initial points $\{p_0^{(i)}\}$ converging to p^* such that the path-dependent Newton iterate started at any $p_0^{(i)}$ does not converge quadratically to p^* . For each $p_0^{(i)}$, the resulting path-dependent Newton iterate is a whimsical Newton method with respect to Σ . From Theorem 12, there exists a neighbourhood of p^* such that any whimsical Newton iterate with respect to Σ that starts within this neighbourhood will converge quadratically to p^* , a contradiction. \square

The motivation given earlier for introducing path-dependent methods was that it is easier and more natural to construct parametrisations locally then extend path-wise than to construct parametrisations globally because the latter requires the local parametrisations to fit together globally. Another use for path-dependent methods is to give the algorithm memory. This leads into the study of general techniques for extending conjugate gradient and other such methods to manifolds, a topic outside the scope of the present paper.

7. Re-Centring and Other Parametrisation Construction Techniques

Various strategies exist for choosing parametrisations that satisfy C1–C2 in Section 5. If the class of cost functions of interest is known beforehand then this knowledge should inform the choice of parametrisation; see Section 4.3. There is also interest in choosing relatively simple parametrisations leading to generic algorithms designed without regard to any particular class of cost functions. A basic idea for how to do this is introduced in Section 7.2 and generalised in subsequent sections. It is called re-centring and exploits the existence of a local diffeomorphism between any two parts of a manifold. *This changes the focus from devising parametrisations to devising transformations.* If M were a sphere, for example, then instead of producing a sequence of points p_1, p_2, \dots on M converging to a critical point p^* , the manifold M , along with the cost function f , can be rotated at each step to bring p_k to the North pole, until eventually the critical point p^* is brought to the North pole. Since each Newton step is always taken from the North pole, its design is simplified.

7.1. Submersions and Fibre Bundles

It may happen that the manifold N for which a parametrisation is sought is the image of a smooth function $g: M \rightarrow N$ where M is simpler to parametrise. For example, M might be a matrix Lie group and N a homogeneous space. Since a cost function $f: N \rightarrow \mathbb{R}$ pulls back to a cost function $f \circ g$ on M , an iterative scheme on M should induce an iterative scheme on N . Simply pulling f back is not recommended if $\dim M > \dim N$ because the final algorithmic complexity might increase and non-degenerate critical points of f can become degenerate critical points of $f \circ g$.

An alternative is to endeavour to “push forwards” the parametrisation on M . Let $g_\star: TM \rightarrow TN$ be the induced push-forward of $g: M \rightarrow N$ and let $p \in M$ be such that g is a submersion at p . Then T_pM splits into a vertical component V_p and a non-unique horizontal component H_p , that is, there exists a subspace $H_p \subset T_pM$ such that $T_pM = V_p \oplus H_p$ where $V_p = \{v \in T_pM \mid g_\star(v) = 0\}$. Since g is a submersion at p , g_\star induces a linear isomorphism from H_p to $T_{g(p)}N$, denoted $g_\star|_{H_p}$. In particular, a parametrisation $\phi_p: T_pM \rightarrow M$ can be used to form the parametrisation $\tilde{\phi}_{g(p)} = g \circ \phi_p \circ (g_\star|_{H_p})^{-1}$ from $T_{g(p)}N$ into N .

Take Σ_q to be the set of all parametrisation pairs $(\tilde{\phi}_q, \tilde{\psi}_q)$ coming from parametrisations $\phi_p, \psi_p: T_pM \rightarrow M$ where $q = g(p)$ and g is a submersion at p . Provided every point $q \in N$ has at least one preimage p such that g is a

submersion at p , a whimsical or path-dependent Newton method is well-defined; see Section 6. Assuming ϕ and ψ satisfy C1–C2 in Section 5, it is not necessarily the case though that the resulting Σ will satisfy E1–E2 in Section 6. There are essentially two ways for E1–E2 to fail to hold. Visually, the first is if the angle (with respect to a Riemannian metric placed on M) between H_p and V_p can approach zero as p varies. The second is if $g^{-1}(q)$ is unbounded and ϕ_p for $p \in g^{-1}(q)$ gets arbitrarily ill-behaved as p goes to infinity.

If $g: M \rightarrow N$ happens to be a compact fibre bundle and an Ehresmann connection is chosen, thereby determining the horizontal bundle, then a compactness argument can be made for E1–E2 to hold.

A more general approach is to limit the number of parametrisation pairs in Σ_q by judiciously choosing which preimages $p \in g^{-1}(q)$ to use. For example, if there exist a locally finite open cover $\{U_\gamma\}$ of N and smooth functions $h_\gamma: U_\gamma \rightarrow M$ such that each $g \circ h_\gamma$ is the identity map, then Σ_q need only contain the finite number of parametrisation pairs coming from the preimages $p = h_\gamma(q)$ for those γ for which $q \in U_\gamma$, and C1–C2 will imply E1–E2.

7.2. Re-centring via a Group Action

If a Lie group G acts transitively on M then a generalised Newton method can be devised by continually re-centring the cost function about a distinguished point. Precisely, fix $\bar{p} \in M$, choose parametrisations $\phi_{\bar{p}}, \psi_{\bar{p}}: T_{\bar{p}}M \rightarrow M$ and define $E_f(\bar{p})$ as in (36). The group action $p \mapsto g \cdot p$ allows $E_f(p)$ to be defined for $p \neq \bar{p}$ by $E_f(p) = h(p) \cdot E_{\tilde{f}}(\bar{p})$ where $h: M \rightarrow G$ is an arbitrary function satisfying $h(p) \cdot \bar{p} = p$, and $\tilde{f}(q) = f(h(p) \cdot q)$ is the re-centred cost function.

If no obvious rule for choosing h comes to mind, the framework of Section 6 can be used. Precisely, define $\Sigma_p = \{(\phi_g, \psi_g) \mid g \in G, g \cdot \bar{p} = p\}$ where $\phi_g = \theta_g \circ \phi_{\bar{p}} \circ Q_g^{-1}$, $\psi_g = \theta_g \circ \psi_{\bar{p}} \circ Q_g^{-1}$, $\theta_g(p) = g \cdot p$ and Q_g is the restriction of $(\theta_g)_*$ to $T_{\bar{p}}M$. Here, $(\theta_g)_*: TM \rightarrow TM$ is the push-forward of $\theta_g: M \rightarrow M$. The parametrisation pair (ϕ_g, ψ_g) yield the same Newton step as before: $\psi_g \circ N_{f \circ \phi_g}(0_{g \cdot \bar{p}}) = g \cdot E_{\tilde{f}}(\bar{p})$ where $\tilde{f}(q) = f(g \cdot q)$. Assume $\phi_{\bar{p}}$ and $\psi_{\bar{p}}$ are sensible (e.g., \mathcal{C}^2 -smooth). Provided the first and second order derivatives of θ_g are bounded in an appropriate sense then E0–E2 will hold; see Lemma 35. This will be the case if G is compact, for example. Otherwise, the number of parametrisations in each Σ_p can be limited with the aid of a finite open cover of M on which local sections are defined; see the end of Section 7.1.

7.3. Re-centring via Affine Transformations

Re-centring can be applied to a manifold M embedded in Euclidean space by using affine transformations of Euclidean space to bring any point $p \in M$ of interest to the origin in such a way that the transformed version of M is a graph of a function in a neighbourhood of the origin. Therefore, a rule for parametrising a graph of a function induces a rule for parametrising M . This idea will be used in Section 7.4. The present section focuses on the reverse direction: determine if a parametrisation pair (ϕ, ψ) satisfies C1–C2 in Section 5 by studying the corresponding re-centred parametrisations.

Denote by \mathfrak{H}_δ the space of \mathcal{C}^2 -smooth functions $h: B_n(0; \delta) \rightarrow \mathbb{R}^k$ satisfying $h(0) = 0$, $Dh(0) = 0$ and $\sup_{t \in B_n(0; \delta)} \|D^2h(t)\| < \infty$. (Recall B_n is an open ball in \mathbb{R}^n .) Associate with any $h \in \mathfrak{H}_\delta$ the manifold

$$\tilde{M}_{(h, \delta)} = \{(t, h(t)) \mid t \in B_n(0; \delta)\} \subset \mathbb{R}^{n+k}. \quad (39)$$

A parametrisation $\tilde{\phi}_{(0,0)}: T_{(0,0)}\tilde{M}_{(h, \delta)} \rightarrow \tilde{M}_{(h, \delta)}$ can be represented by a function $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ taking x to the point $(\pi(x), h \circ \pi(x))$ on $\tilde{M}_{(h, \delta)}$. Here, x represents the point $(x, 0)$ on the affine tangent space at the origin of $\tilde{M}_{(h, \delta)}$. An exemplar is using Euclidean projection from $(x, 0)$ to $\tilde{M}_{(h, \delta)}$ to define π , so that π satisfies

$$\|(\pi(x), h \circ \pi(x)) - (x, 0)\| = \min_{t \in B(0; \delta)} \|(t, h(t)) - (x, 0)\| \quad (40)$$

whenever the minimum exists. This will be studied in Section 7.4.

It is not important for π to be defined uniquely by a rule, nor for π to be defined on the whole of \mathbb{R}^n . Essentially, it is merely required that any choice of π is defined on a sufficiently small domain $B_n(0; \rho)$ and satisfies $\|D^2\pi(0)\| \leq \alpha$ and $\|\pi(x) - x\| \leq \beta\|x\|^2$ for $x \in B_n(0; \rho)$, where the constants ρ , α and β depend on h and δ in a way that ensures they remain uniformly bounded if h and δ are perturbed; see Section 6. This is now made precise.

A parametrisation $\phi: TM \rightarrow M$ of an n -dimensional embedded submanifold $M \subset \mathbb{R}^{n+k}$ is said to satisfy condition **D** if D1–D6 below are satisfied. Central to this condition is a class of functions π obtained from ϕ as follows. Associate to each $x \in M \subset \mathbb{R}^{n+k}$ a rotation R_x of \mathbb{R}^{n+k} sending V_xM to $\mathbb{R}^n \times \{0\}$. (Here, the notation of Section Appendix C.3 is being used, but with i and i_* omitted because M is being treated as an actual subset of \mathbb{R}^{n+k} .) Define the translation $Q_x(q) = q - x$. For any pair (h, δ) for which $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h, \delta)}$, meaning there exists a set U open in \mathbb{R}^{n+k} such that $R_x Q_x(M) \cap U = \tilde{M}_{(h, \delta)}$, a function $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ can be defined implicitly by

$$(\pi(y), h \circ \pi(y)) = R_x Q_x \hat{\phi}(x, R_x^{-1} Jy) \quad (41)$$

where $\hat{\phi} = \phi \circ \tau_I^{-1}$ is the representation of ϕ in local coordinates with respect to the identity chart (on \mathbb{R}^{n+k} ; refer to Section Appendix C.3) and $J: \mathbb{R}^n \rightarrow \mathbb{R}^{n+k}$ sends y to $J(y) = (y, 0)$. To emphasise, π is only defined at points $y \in \mathbb{R}^n$ for which (41) holds. Note too that π , which depends on the triple x , h and δ , is merely ϕ_x , the restriction of ϕ to T_xM , written in a canonical form (albeit depending on the choice of rotation R_x).

Condition D requires there to exist functions $\alpha, \beta, \rho: (0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$ such that for all $x \in M$, $\delta > 0$ and $h \in \mathfrak{H}_\delta$ for which $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h, \delta)}$ and for $K = \sup_{t \in B_n(0; \delta)} \|D^2h(t)\|$:

D1 The domain of definition of π in (41) includes $B_n(0; \rho(\delta, K))$;

D2 $\|D^2\pi(0)\| \leq \alpha(\delta, K)$;

D3 $\|\pi(y) - y\| \leq \beta(\delta, K) \|y\|^2$ for $y \in B_n(0; \rho(\delta, K))$.

It is also required that for all $\delta > 0$ and all $\bar{K} \in [0, \infty)$:

D4 $\sup_{K \in [0, \bar{K}]} \alpha(\delta, K) < \infty$;

D5 $\sup_{K \in [0, \bar{K}]} \beta(\delta, K) < \infty$;

D6 $\inf_{K \in [0, \bar{K}]} \rho(\delta, K) > 0$.

A sufficient condition for D4–6 to hold is for α and β to be upper semi-continuous in K , and ρ lower semi-continuous in K .

Proposition 14. *Let $\phi: TM \rightarrow M$ be a parametrisation of an n -dimensional \mathcal{C}^2 -smooth embedded submanifold $M \subset \mathbb{R}^{n+k}$. If ϕ satisfies condition D described above then ϕ and $\psi = \phi$ satisfy C1–C2 in Section 5. (If ϕ satisfies D1, D2, D4 and, instead of D3 and D6, the weaker conditions that $\pi(0) = 0$, $D\pi(0) = I$ and $\rho(\delta, K) > 0$, then ϕ satisfies C1. If ϕ satisfies D1, D3, D5 and D6 then $\psi = \phi$ satisfies C2.)*

Proof. When convenient, elements of \mathbb{R}^{n+k} are written as $(u, v) \in \mathbb{R}^n \times \mathbb{R}^k$, with projections P_1 and P_2 sending (u, v) to u and v respectively. Fix a point $z \in M \subset \mathbb{R}^{n+k}$. Let $\delta > 0$ and $h \in \mathfrak{H}_{6\delta}$ be such that $\tilde{M}_{(h, 6\delta)} = R_z Q_z(M) \cap U$ for some open set $U \subset \mathbb{R}^{n+k}$. Let $K = \sup_{t \in B(0; 6\delta)} \|D^2 h(t)\|$. By shrinking δ if necessary, it is assumed without loss of generality that $\delta < 1/(4K)$ and $\|Dh(t)\| < 1/4$ for $t \in B(0; 6\delta)$.

Choose an arbitrary $t \in B(0; \delta)$. Define $x = Q_z^{-1} R_z^{-1}(t, h(t))$ and $f_t(\tau) = P_1 R_x R_z^{-1}(\tau, h(t + \tau) - h(t))$ for $\tau \in B(-t; 6\delta) \supset B(0; 5\delta)$. It will be shown that $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h_t, \delta)}$ where $h_t(u) = P_2 R_x R_z^{-1}(f_t^{-1}(u), h(t + f_t^{-1}(u)) - h(t))$.

First, bounds on f_t and its derivatives $Df_t(\tau) \cdot \xi = P_1 R_x R_z^{-1}(\xi, Dh(t + \tau) \cdot \xi)$ and $D^2 f_t(\tau) \cdot \xi^2 = P_1 R_x R_z^{-1}(0, D^2 h(t + \tau) \cdot \xi^2)$ are obtained. Importantly, $P_2 R_x R_z^{-1}(\xi, Dh(t) \cdot \xi) = 0$ because $R_z^{-1}(\xi, Dh(t) \cdot \xi)$ lies in $V_x M$, the latter a consequence of the curve $\gamma(\theta) = Q_z^{-1} R_z^{-1}(t + \theta\xi, h(t + \theta\xi))$ lying on M and having $\gamma'(0) = R_z^{-1}(\xi, Dh(t) \cdot \xi)$. Therefore, $\|Df_t(0) \cdot \xi\| = \|(\xi, Dh(t) \cdot \xi)\| \geq \|\xi\|$. Similarly, $\|f_t(\tau)\| \geq \|\tau\| - \|h(t + \tau) - h(t) - Dh(t) \cdot \tau\| \geq \|\tau\| - \frac{1}{2}K\|\tau\|^2$ (see Lemma 33). For $\tau_1, \tau_2 \in B(-t; 6\delta)$, $\|Df_t(\tau_1) - Df_t(\tau_2)\| \leq \|Dh(t + \tau_1) - Dh(t + \tau_2)\| < \frac{1}{2}$. Thus $\|(Df_t(\tau))^{-1}\| < 2$ because, for $\|\xi\| > 0$, $\|Df_t(\tau) \cdot \xi\| \geq \|Df_t(0) \cdot \xi\| - \|(Df_t(\tau) - Df_t(0)) \cdot \xi\| > \frac{1}{2}\|\xi\|$. Lemma 1 of [13, Chapter 16] implies f_t is injective on $B(-t; 6\delta)$, and, since Df_t is invertible, the inverse function theorem implies f_t is a \mathcal{C}^2 -diffeomorphism from $B(-t; 6\delta)$ onto its image. Furthermore, since $\|\tau\| = 4\delta$ implies $\|f_t(\tau) - f_t(0)\| \geq 4\delta - \frac{1}{2}K(4\delta)^2 > 2\delta$, the proof of Lemma 2 of [13, Chapter 16] implies $B(0; \delta) \subset f_t(B(0; 4\delta))$. Thus, $h_t(u)$ is a well-defined \mathcal{C}^2 -smooth function on $B(0; \delta)$. Finally, note $\|D^2 f_t(\tau)\| \leq \|D^2 h(t + \tau)\| \leq K$.

Next it is shown that $\tilde{M}_{(h_t, \delta)} = R_x R_z^{-1} Q_{(t, h(t))}(\tilde{M}_{(h, 6\delta)}) \cap P_1^{-1}(B(0; \delta))$. Indeed, if $(u, h_t(u)) \in \tilde{M}_{(h_t, \delta)}$ then $\tau = f_t^{-1}(u)$ is well-defined and $(t + \tau, h(t + \tau)) \in \tilde{M}_{(h, 6\delta)}$ is such that $(u, h_t(u)) = R_x R_z^{-1} Q_{(t, h(t))}(t + \tau, h(t + \tau))$. Conversely, an arbitrary element p of $R_x R_z^{-1} Q_{(t, h(t))}(\tilde{M}_{(h, 6\delta)})$ is of the form $p =$

$R_x R_z^{-1}(\tau, h(t + \tau) - h(t))$ for some $\tau \in B(-t; 6\delta)$. Since $P_1(p) = f_t(\tau)$, if $p \in P_1^{-1}(B(0; \delta))$ then $u = f_t(\tau) \in B(0; \delta)$ and $p = (u, h_t(u)) \in \tilde{M}_{(h_t, \delta)}$. Because $(t, h(t)) = R_z Q_z(x)$, it follows that $R_x R_z^{-1} Q_{(t, h(t))} R_z Q_z = R_x Q_x$. Therefore, $\tilde{M}_{(h_t, \delta)} = R_x Q_x(M) \cap R_x Q_x Q_z^{-1} R_z^{-1}(U) \cap P_1^{-1}(B(0; \delta))$, proving $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h_t, \delta)}$.

To show $h_t \in \mathfrak{H}_\delta$, first note $h_t(0) = 0$ and $Dh_t(0) = 0$, the latter a consequence of $P_2 R_x R_z^{-1}(\xi, Dh(t) \cdot \xi) = 0$ for all ξ . It remains to bound $\|D^2 h_t(u)\|$. For an arbitrary $u \in B(0; \delta)$, let $\tau = f_t^{-1}(u) \in B(0; 4\delta)$. Then $D^2 h_t(u) \cdot \xi^2 = P_2 R_x R_z^{-1}(D^2 f_t^{-1}(u) \cdot \xi^2, D^2 h(t + \tau) \cdot (Df_t^{-1}(u) \cdot \xi)^2 + Dh(t + \tau) D^2 f_t^{-1}(u) \cdot \xi^2)$. Now, $Df_t^{-1}(u) = (Df_t(\tau))^{-1}$ and $D^2 f_t^{-1}(u) \cdot \xi^2 = -(Df_t(\tau))^{-1} D^2 f_t(\tau) \cdot ((Df_t(\tau))^{-1} \cdot \xi)^2$. Applying the earlier bounds shows $\|D^2 h_t(u)\| \leq \bar{K}$ where $\bar{K} = \sqrt{(8K)^2 + (4K + 2K)^2} = 10K$.

With the values of δ and \bar{K} as above, and referring to D4–D6, define $\bar{\alpha} = \sup_{K \in [0, \bar{K}]} \alpha(\delta, K)$, $\bar{\beta} = \sup_{K \in [0, \bar{K}]} \beta(\delta, K)$, and $\bar{\rho} = \inf_{K \in [0, \bar{K}]} \rho(\delta, K)$. Shrink $\bar{\rho}$ if necessary to ensure $\bar{\rho} < \delta$ and $B_{n+k}(0; 2\bar{\rho}) \subset U$. Thus, for any $x \in M \cap B_{n+k}(z; \bar{\rho})$ there exists a $t \in B(0; 6\delta)$ such that $x = Q_z^{-1} R_z^{-1}(t, h(t))$, in which case $\|t\| \leq \|R_z^{-1}(t, h(t))\| = \|x - z\| < \bar{\rho} < \delta$. In particular then, $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h_t, \delta)}$. For the particular triple x, h_t and δ , define π as in (41). From D1–D3 it follows that π is defined on $B(0; \bar{\rho})$ and satisfies $\|D^2 \pi(0)\| \leq \bar{\alpha}$ and $\|\pi(\tilde{y}) - \tilde{y}\| \leq \bar{\beta} \|\tilde{y}\|^2$ for $\tilde{y} \in B(0; \bar{\rho})$.

The proof is completed by showing $\hat{\phi}$ satisfies the assumptions in Lemma 43. Let $y \in B(0; \bar{\rho}) \cap V_x M$ be arbitrary. Define $\tilde{y} = P_1 R_x y$ and note $R_x y = (\tilde{y}, 0)$. From D1 and (41), since $\|\tilde{y}\| = \|y\| < \bar{\rho}$, $\hat{\phi}_x(y) = Q_x^{-1} R_x^{-1}(\pi(\tilde{y}), h_t \circ \pi(\tilde{y}))$. Then $\hat{\phi}_x(y)$ lies in M because $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h_t, \delta)}$. That $\hat{\phi}_x(0) = 0$ follows from $\pi(0) = 0$ (D3) and $h_t(0) = 0$. The facts $\pi(0) = 0$, $D\pi(0) = I$ (D3) and $Dh_t(0) = 0$ imply $D\hat{\phi}_x(0) \cdot y = y$ and $D^2 \hat{\phi}_x(0) \cdot y^2 = R_x^{-1}(D^2 \pi(0) \cdot \tilde{y}^2, D^2 h_t(0) \cdot \tilde{y}^2)$. Thus $\|D^2 \hat{\phi}_x(0) \cdot y^2\| \leq \sqrt{\bar{\alpha}^2 + \bar{K}^2} \|y\|^2$. Lemma 33 implies $\|h_t(u)\| \leq \frac{1}{2} \bar{K} \|u\|^2$. Therefore, $\|h_t \circ \pi(\tilde{y})\| \leq \frac{1}{2} \bar{K} (\|\pi(\tilde{y}) - \tilde{y}\|^2 + \|\tilde{y}\|^2) \leq \frac{1}{2} \bar{K} (1 + \bar{\beta}^2 \bar{\rho}^2) \|\tilde{y}\|^2$. Thus, $\|\hat{\phi}_x(y) - x - y\| = \|(\pi(\tilde{y}) - \tilde{y}, h_t \circ \pi(\tilde{y}))\| \leq \sqrt{\bar{\beta}^2 + (1/4)\bar{K}^2(1 + \bar{\beta}^2 \bar{\rho}^2)} \|y\|^2$. \square

7.4. Local and Global Projections

For an embedded manifold $M \subset \mathbb{R}^{n+k}$, it was suggested in [15] that projection from the affine tangent plane to the manifold could be used to define parametrisations. The utility of Proposition 14 is illustrated by proving such parametrisations satisfy C1–C2. Only ordinary calculus is required as the differential geometric framework is hidden behind Proposition 14. Note the proof works at the generality of \mathcal{C}^2 -smooth manifolds and is thus not based on a smoothness argument (Lemma 39).

Lemma 15. *Let $h: B_n(0; \delta) \rightarrow \mathbb{R}^k$ be a \mathcal{C}^2 -smooth map with $h(0) = 0$ and $Dh(0) = 0$. Let $K = \sup_{x \in B(0; \delta)} \|D^2 h(x)\|$ and $\rho = \frac{1}{2} \min\{\delta, \sqrt{2}/(3K)\}$. There exists a \mathcal{C}^1 -smooth map $\pi: B_n(0; \rho) \rightarrow B_n(0; 2\rho)$ such that, for any $x \in B(0; \rho)$, $(\pi(x), h \circ \pi(x))$ is the unique point on the manifold $\{(t, h(t)) \mid t \in B(0; \delta)\}$*

closest to $(x, 0)$. Moreover, no point is closest to $(x, 0)$ on the smaller manifold $\{(t, h(t)) \mid \|t\| < \|\pi(x)\|\}$. The map π satisfies $\pi(0) = 0$, $D\pi(0) = I$, $D^2\pi(0) = 0$ and $\|\pi(x) - x\| \leq \frac{1}{2}K\|x\|^2$.

Proof. Define $f(x) = x + g(x)$ and $g(x) = (Dh(x))^T h(x)$ where superscript T denotes adjoint. Then $(f(x), 0)$ is the unique point of intersection of the affine plane normal to the manifold at $(x, h(x))$ with the plane $\mathbb{R}^n \times 0$. Thus if π exists it must satisfy $f(\pi(x)) = x$.

From Lemma 33, $\|h(x)\| \leq (K/2)\|x\|^2$ and $\|Dh(x)\| \leq K\|x\|$ for $x \in B(0; \delta)$. Therefore, $\|g(x)\| \leq (1/2)K^2\|x\|^3$, $\|Dg(x)\| \leq (3/2)K^2\|x\|^2$, $\|Df(x) \cdot \xi\| \geq (1 - (3/2)K^2\|x\|^2)\|\xi\|$ and $\|Df(x) - Df(y)\| \leq (3/2)K^2(\|x\|^2 + \|y\|^2)$ for $x, y \in B(0; \delta)$. The latter implies $D^2f(0) = 0$. If $x, y \in B(0; 2\rho)$ then $\|(Df(x))^{-1}\| < 3/2$ and $\|Df(x) - Df(y)\| < 2/3$ so Lemma 1 of [13, Chapter 16] and the inverse function theorem imply f restricted to $B(0; 2\rho)$ is a \mathcal{C}^1 -diffeomorphism.

Let $x \in B(0; \rho)$. Since the manifold includes the origin, a distance $\|x\|$ away from $(x, 0)$, the closest point(s) to $(x, 0)$ on the original manifold are the same as the closest point(s) on the smaller manifold $\{(t, h(t)) \mid t \in \overline{B}(0; 2\|x\|)\}$. The latter manifold is compact and hence a closest point exists. Uniqueness follows from f being injective on $B(0; 2\rho) \supset \overline{B}(0; 2\|x\|)$; any closest point $(t, h(t))$ must satisfy $f(t) = x$. Since this is a local condition, it also means no point is closest to $(x, 0)$ on $\{(t, h(t)) \mid \|t\| < \|\pi(x)\|\}$. Note that $\|\pi(x)\| \leq 2\|x\|$.

The geometric bound $\|\pi(x) - x\| \leq \|h(x)\|$ implies $\|\pi(x) - x\| \leq (K/2)\|x\|^2$, so $\pi(0) = 0$ and $D\pi(0) = I$. As f is a \mathcal{C}^1 -diffeomorphism on $B(0; 2\rho)$, $\pi(x) = f^{-1}(x)$ for $x \in B(0; \rho)$ is \mathcal{C}^1 -smooth and $D\pi = (Df \circ \pi)^{-1}$. For $\epsilon > 0$, choose $\delta > 0$ such that $\|y\| < 2\delta$ implies $\|y\| \leq 2\|f(y)\|$ and $\|I - Df(y)\| < \epsilon\|y\| < \frac{1}{2}$. (This is possible because $D^2f(0) = 0$.) Then for $\|x\| < \min\{\rho, \delta\}$ and $y = \pi(x)$, $\|D\pi(x) - I\| = \|(I - (I - Df(y)))^{-1} - I\| \leq (1 - \|I - Df(y)\|)^{-1} - 1 < 2\epsilon\|y\| \leq 4\epsilon\|f(y)\| = 4\epsilon\|x\|$, proving $D^2\pi(0) = 0$. \square

It is mentioned tangentially that the cubic bound $\|\pi(x) - x\| = \|f(\pi(x)) - \pi(x)\| \leq 4K^2\|x\|^3$ is readily obtainable from the above proof.

Given a rule such as (40), a parametrisation of a manifold can be obtained by using (41) in reverse. The only technicality is the choice of neighbourhood size δ to use for each point on the manifold. The actual choice is generally not important provided a positive uniform lower bound exists on any compact neighbourhood. In fact, as is the case in the following lemma, the choice may depend on y as well as x in (41).

Proposition 16. *Let $M \subset \mathbb{R}^{n+k}$ be an embedded \mathcal{C}^2 -manifold of dimension n . Adopting the notation of Section 7.3, associate with each $x \in M$ and $y \in \mathbb{R}^n$ a $\delta_{xy} > 0$, an $h_{xy} \in \mathfrak{H}_{\delta_{xy}}$ and a rotation R_x such that: 1) $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h_{xy}, \delta_{xy})}$; and, 2) if no point of $\tilde{M}_{(h_{xy}, \delta_{xy})}$ is closest to $(y, 0)$ then the same is true for any admissible choice of δ_{xy} . Referring to (41) and (40), if $(t, h_{xy}(t)) \in \tilde{M}_{(h_{xy}, \delta_{xy})}$ is the unique closest point to $(y, 0)$ then set $\hat{\phi}(x, R_x^{-1}Jy) = Q_x^{-1}R_x^{-1}(t, h_{xy}(t))$. Otherwise, if the closest point does not exist or is not unique, let $\hat{\phi}(x, R_x^{-1}Jy)$ be an arbitrary element of M . Then the parametrisation $\phi: TM \rightarrow M$, $\phi = \hat{\phi} \circ \tau_I$, satisfies condition D.*

Proof. The $k = 0$ case is straightforward so assume $k > 0$. Define the functions $\rho(\delta, K) = \frac{1}{2} \min\{\delta, \sqrt{2}/(3K)\}$, $\alpha(\delta, K) = 0$ and $\beta(\delta, K) = K/2$; they satisfy D4–D6. Next, choose $x \in M$, $\delta > 0$ and $h \in \mathfrak{H}_\delta$ such that $R_x Q_x(M)$ locally looks like $\tilde{M}_{(h,\delta)}$. Let $K = \sup_{t \in B(0;\delta)} \|D^2 h(t)\|$. Then for $y \in B(0; \rho(\delta, K))$ define $\pi(y)$ as in Lemma 15; the unique closest point to $(y, 0)$ on $\tilde{M}_{(h,\delta)}$ is $(\pi(y), h \circ \pi(y))$. This must therefore correspond with the π in (41). That D1–D3 hold follows immediately from Lemma 15. \square

To assist in interpreting Proposition 16, consider how the projection from $(y, 0)$ onto $\tilde{M}_{(h,\delta)}$ changes with δ . If δ is too small then no point is necessarily closest because $\tilde{M}_{(h,\delta)}$ is not compact. As δ increases the closest point may change as more candidates become available. The advantage of Proposition 16 in practice is it allows parametrisations to be defined using only local minima of the Euclidean distance function rather than insisting on global minima. Note too that projecting onto $\tilde{M}_{(h,\delta)}$ is different from projecting onto M because M may curve around and come close to touching itself.

Proposition 17. *Let $M \subset \mathbb{R}^{n+k}$ be an embedded n -dimensional C^2 -manifold. Let $a: TM \rightarrow \mathbb{R}^{n+k}$ be the map taking a tangent vector to its equivalent point on the affine tangent plane. Let $\phi: TM \rightarrow M$ be any map with the property that $\|\phi(v_p) - a(v_p)\| = \min_{q \in M} \|q - a(v_p)\|$ whenever the minimum exists, where the norm is the Euclidean norm on \mathbb{R}^{n+k} . Then ϕ and $\psi = \phi$ satisfy C1 and C2 of Section 5.*

Proof. The $k = 0$ case is straightforward so assume $k > 0$. For $z \in M$, define $\rho(z) = \sup\{\rho \mid \exists h \in \mathfrak{H}_\rho, \exists \text{ an open } U \supset B_{n+k}(0; \rho), R_z Q_z(M) \cap U = \tilde{M}_{(h,\rho)}\}$; see Section 7.3 for notation. Let $K \subset M$ be a compact set and assume to the contrary there exists a convergent sequence $z_i \rightarrow \bar{z}$ in K with $\rho(z_i) \rightarrow 0$. It follows from the proof of Proposition 14 that at \bar{z} there exist a $\delta > 0$ and an $h \in \mathfrak{H}_{6\delta}$ such that (by shrinking δ if necessary) $\tilde{M}_{(h,6\delta)} = R_{\bar{z}} Q_{\bar{z}}(M) \cap U$, $U \supset B_{n+k}(0; 6\delta)$ and for any $z \in B(\bar{z}; \delta) \cap M$, there exists an $h_z \in \mathfrak{H}_\delta$ such that $\tilde{M}_{(h_z,\delta)} = R_z Q_z(M) \cap R_z Q_z Q_{\bar{z}}^{-1} R_{\bar{z}}^{-1}(U) \cap P_1^{-1}(B_n(0; \delta))$. Since $B_{n+k}(0; \delta) \subset R_z Q_z Q_{\bar{z}}^{-1} R_{\bar{z}}^{-1}(B_{n+k}(0; 6\delta))$ it follows that $\rho(z) > \rho$ for $\|z - \bar{z}\| < \rho$, a contradiction. Thus $\inf_{z \in K} \rho(z) > 0$.

For $x \in M$ and $y \in B_n(0; \rho(x)/4)$ define $\delta_{xy} = \rho(x)/2$ and $\hat{\phi}(x, R_x^{-1} Jy) = \phi \circ \tau_I^{-1}(x, R_x^{-1} Jy)$. The closest point to $(y, 0)$ on $R_x Q_x(M)$ must be contained in $B_{n+k}(0; \rho(x)/2)$ and hence is in the local representation $\tilde{M}_{(h_{xy}, \delta_{xy})}$. It is therefore possible to define δ_{xy} and $\hat{\phi}(x, R_x^{-1} Jy)$ for $y \notin B_n(0; \rho(x)/4)$ so that $\hat{\phi}$ satisfies the conditions of Proposition 16 (and hence by Proposition 14 the corresponding parametrisation satisfies C1 and C2). By Lemma 42, because $\phi(v_x) = \hat{\phi} \circ \tau_I(v_x)$ whenever $\|v_x\| < \rho(x)/4$, ϕ and ψ satisfy C1 and C2. \square

Since any manifold can be embedded in \mathbb{R}^{n+k} for sufficiently large k , Proposition 17 guarantees the existence of parametrisations satisfying C1–C2.

7.5. Discussion

To the best of our knowledge, all Newton methods on finite-dimensional manifolds in the literature can be rewritten as (36) where the parametrisations ϕ and ψ are smooth. Theorem 11 and Lemma 39 together imply that such Newton methods have local quadratic convergence. As a specific example, the original Riemannian Newton method in [11] uses the Riemannian exponential map for the parametrisations ϕ and ψ ; see (2). It is a standard result that if M is \mathcal{C}^4 -smooth then Exp is \mathcal{C}^2 -smooth on a neighbourhood of the zero section, and moreover, $\text{Exp}_p(0) = p$ and $D\text{Exp}_p(0) = I$. Therefore, the Riemannian Newton method (2) has local quadratic convergence by Remark 40.

The article [15] introduced Newton methods on (real and complex) Grassmann and Stiefel manifolds, with parametrisations chosen to be global projections from Euclidean space onto the affine tangent planes of the Stiefel manifold, and an analogous choice made for the Grassmann manifold by treating it as a quotient space of the Stiefel manifold. Local quadratic convergence follows from Proposition 17 and Section 7.1.

When sufficient smoothness is not present for Lemma 39 to be applicable, the proofs in Section 7 demonstrate that essentially all the effort goes into obtaining *uniform* bounds. Condition D in Section 7.3 is one illustration of this.

The conjecture made in Section 4.1 applies to Newton methods on manifolds too. It is difficult to see how any iterative scheme can fail to be of the form (36) if it uses only the information in the 2-jet of f about the current point to converge locally quadratically to a non-degenerate critical point for a sufficiently rich class of functions f .

An advantage of expressing an algorithm in the form (36) is that it gives the algorithm the following heuristic interpretation: at each step, the parametrisation ϕ endeavours to make $f \circ \phi$ look as quadratic as possible, while ψ endeavours to map the result back to the manifold as cheaply as possible; see Section 4.3. Additionally, the fundamental idea of re-centring can further simplify matters.

Finally, it is remarked that affine connections and parallel transport can be used to construct parametrisations. This can be understood in terms of the classical notion of development in differential geometry; a manifold M can be rolled along an affine space without slipping. This is a representative example of the re-centring technique in Section 7.3.

8. Iterates Computing Coordinate Independent Properties

This section studies how iteration functions besides the Newton iteration function can be lifted from Euclidean space to manifolds. This necessitates introducing a rudimentary theory of iterative methods computing coordinate independent properties. It also studies further the generalised Newton method at a grass-roots level.

First, the concept of converging to an identifiable point of f needs defining. Example 20 may prove illuminative.

Definition 18. Assign to each $f: \text{Dom } f \subset \mathbb{R}^n \rightarrow \mathbb{R}$ a subset $P_f \subset \text{Int Dom } f$ of the interior of the domain of f . The *property* $P = \{P_f\}$ is \mathcal{C}^k -coordinate independent if $x \in P_f$ implies $\phi^{-1}(x) \in P_{f \circ \phi}$ for every \mathcal{C}^k -diffeomorphism ϕ of open sets in \mathbb{R}^n with x in the image of ϕ .

Henceforth, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ will mean $f: \text{Dom } f \subset \mathbb{R}^n \rightarrow \mathbb{R}$ with an implicit requirement that a particular point be in the domain of f whenever necessary. For example, f being \mathcal{C}^k -smooth at x implicitly requires $x \in \text{Dom } f$.

Two functions $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ are k -jet equivalent at $p \in \mathbb{R}^n$ if f and g are \mathcal{C}^k -smooth in a neighbourhood of p and $f(p) = g(p)$, $Df(p) = Dg(p)$, \dots , $D^k f(p) = D^k g(p)$.

Definition 19. A k th-order iterative method is the assignment of an iteration function $N_f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ to each $f: \mathbb{R}^n \rightarrow \mathbb{R}$ where $N_f(p) = N_g(p)$ whenever f and g are k -jet equivalent at p . An iterative method *computes the property* $P = \{P_f\}$ with rate K if, for any given f and $x^* \in P_f$, the iterate N_f converges locally to x^* with rate K .

Example 20. Let P_f be the set of points x such that f is \mathcal{C}^3 -smooth in a neighbourhood of x , and x is a non-degenerate critical point of f . Then P is \mathcal{C}^3 -coordinate independent and the Newton iterate N_f in (1) is a 2nd-order iterative method that computes P with rate 2.

Example 21. In Example 20, P_f can be instead the set of points x^* satisfying the conditions in Theorem 1, including (4). Then P is a \mathcal{C}^2 -coordinate independent property (see Remark 6) computed by the Newton iterate (1).

Remark 22. It follows from Definition 18 by using the identity map $\phi: U \rightarrow U$ that if P is a \mathcal{C}^k -coordinate independent property, U is an open subset of \mathbb{R}^n and $g = f|_U$ is the restriction to U of a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ then $P_f \cap U \subset P_g$. The converse is not implied; properties can be “forgotten” as the domain increases. Stricter definitions precluding this are not necessary for what follows.

Properties of functions in \mathbb{R}^n lift to properties of functions on manifolds.

Definition 23. Let M be a manifold with maximal atlas \mathcal{A} of \mathcal{C}^k -smooth charts (U, φ) . Let $P = \{P_f\}$ be a $\mathcal{C}^{k'}$ -coordinate independent property with $k' \leq k$. For any $f: M \rightarrow \mathbb{R}$, define $\bar{P}_f = \{p \in M \mid \exists (U, \varphi) \in \mathcal{A}, p \in U, \varphi(p) \in P_{f \circ \varphi^{-1}}\}$. The elements of \bar{P}_f are said to have property P .

Lemma 24. Let $P = \{P_f\}$ be a \mathcal{C}^k -coordinate independent property with $k \geq 1$. Let N be an iterative method computing P with rate $K > 1$. Fix an $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and $x^* \in P_f$. Let ϕ be a \mathcal{C}^k -diffeomorphism of open subsets of \mathbb{R}^n whose image contains x^* . Then the iteration function

$$\bar{N}_f = \phi \circ N_{f \circ \phi} \circ \phi^{-1} \tag{42}$$

converges locally with rate K to x^* .

Proof. Since $\phi^{-1}(x^*) \in P_{f \circ \phi}$, $N_{f \circ \phi}$ converges locally to $\phi^{-1}(x^*)$ with rate K . Furthermore, ϕ is bi-Lipschitz about $\phi^{-1}(x^*)$ in a suitably small neighbourhood of $\phi^{-1}(x^*)$ because $k \geq 1$. Lemma 26 completes the proof. \square

Lemma 24 suggests the coordinate-adapted viewpoint used to generalise the Newton method to manifolds may prove beneficial in more general contexts. The remainder of this section elicits this idea.

The convergence proofs for the coordinate adapted Newton method in Section 4 and the generalised Newton method in Section 5 relied on essentially just two properties of the Newton iterate: invariance to affine coordinate changes, and a lower bound on the radius of convergence of the Newton iterate $N_{f \circ \phi}$ in terms of the second-order behaviour of ϕ . Here, radius of convergence refers to $\bar{\rho}$ in Section 2.

Affine invariance was exploited partially for convenience — it meant only parametrisations ϕ_p with $\phi_p(0_p) = p$ and $D\phi_p(0_p) = I$ were needed — and partially to allow the Newton iterate (1) to be applied unambiguously to the abstract vector space $T_p M$. The reason for using tangent spaces was again for convenience. It made it easier to exploit smoothness when possible. Section 5.1 discussed this in detail.

Affine invariance was also used for re-centring. In (26), the change of coordinate transformations ϕ_x do not change the point x , that is, $\phi_x(x) = x$. This is perhaps the most natural choice for $\phi_x(x)$ as it does not shift the space unnecessarily. When lifting an iteration function to a manifold, such a choice is no longer possible. The proposed solution was to choose a distinguished point of \mathbb{R}^n , the origin, and always apply the Newton iteration function at this distinguished point; see (36). The invariance of the Newton method to shifts made this inconsequential.

If the iterative method N_f is not shift-invariant then re-centring it at each iteration may alter its behaviour. It is therefore necessary to study the re-centred iterate $x_{k+1} = \theta_{x_k} \circ N_{f \circ \theta_{x_k}}(0)$ where $\theta_x(y) = x + y$. Equivalently, N can be replaced by its re-centred version $\tilde{N}_f(x) = \theta_x \circ N_{f \circ \theta_x}(0)$ which is shift-invariant: $\theta_z \circ \tilde{N}_f \circ \theta_z^{-1}(x) = \tilde{N}_f(x)$. *Henceforth the iterative method N_f is assumed to be shift-invariant.*

If the iterative method N_f is affine-invariant then smooth parametrisations $\phi: TM \rightarrow M$ can be used to lift N_f to manifolds in the same way the Newton method was lifted. Otherwise, parametrisations from $M \times \mathbb{R}^n$ rather than TM need be considered if a global approach is taken. A simpler and more general alternative is to construct parametrisations locally, as in Section 6, leading to path-dependent lifts of N_f .

Henceforth, a local viewpoint is adopted because determining how to make parametrisations constructed locally fit together globally is a topological problem unrelated to local convergence properties and which needs no addressing if a path-dependent lift is adequate.

In one sense, lifting a shift-invariant iterative method to a manifold locally about a point is straightforward.

Proposition 25. *As in Definition 23, let M be a C^k -smooth manifold, let (U, φ) be a chart on M , and let \bar{P} be the lift to M of a $C^{k'}$ -coordinate independent property P , with $k' \leq k$. Let N an iterative method of order at most k that computes P with rate K . Define $E_f(p) = \varphi^{-1} \circ \theta_{\varphi(p)} \circ N_{f \circ \varphi^{-1} \circ \theta_{\varphi(p)}}(0)$ for $p \in U$, where $\theta_x(y) = x + y$. Then E_f computes \bar{P} on U with rate K , meaning for any cost function $f: M \rightarrow \mathbb{R}$ and any $p^* \in U \cap \bar{P}_f$, the iteration function E_f converges locally with rate K to p^* (Definition 28).*

Proof. Let (V, ψ) be such that $\psi(p^*) \in P_{f \circ \psi^{-1}}$. Shrink U if necessary so that $p^* \in U \subset V$. Since P is coordinate invariant, the diffeomorphism $\psi \circ \varphi^{-1}$ can be used to show $\varphi(p^*) \in P_{f \circ \varphi^{-1}}$. By Definition 28, it suffices to study $\tilde{N}(x) = \varphi \circ E_f \circ \varphi^{-1}(x) = \theta_x \circ N_{(f \circ \varphi^{-1}) \circ \theta_x}(0) = \theta_x \circ N_{(f \circ \varphi^{-1}) \circ \theta_x} \circ \theta_x^{-1}(x) = N_{(f \circ \varphi^{-1})}(x)$ where the last equality follows from shift-invariance. Since $x^* = \varphi(p^*) \in P_{f \circ \varphi^{-1}}$ and N computes P , \tilde{N} converges locally to x^* with rate K , as required. \square

Allowing more flexibility than afforded by Proposition 25 is desirable for two reasons: Section 4.3 explained how customised parametrisations can improve performance for certain classes of cost functions, and Section 7 gave techniques for adapting parametrisations to geometric features of the manifold.

The most general way found for lifting a Newton method to a manifold is (36). Furthermore, the use of ψ in (36) is an add-on: if $E_f(p) = \phi_p \circ N_{f \circ \phi_p}(0_p)$ converges then (36) will also converge with the same rate provided ψ is a sufficiently good approximation to ϕ . All that remains then is to understand when $E_f(p) = \phi_p \circ N_{f \circ \phi_p}(0)$ computes \bar{P} with rate K given that N computes P with rate K , as in Proposition 25. Note that here, ϕ is defined on $U \times \mathbb{R}^n$ where U is an open subset of M , and $\phi_p(0) = p$. It is also necessary for ϕ_p to be a local diffeomorphism about 0, that is, a genuine change of coordinates.

As in the proof of Proposition 25, consider $\tilde{N}(x) = \varphi \circ E_f \circ \varphi^{-1}(x)$. Using the fact that $(\varphi \circ \phi_{\varphi^{-1}(x)} \circ \theta_{-x})(x) = x$ for all x , this becomes

$$\tilde{N}(x) = \sigma_x \circ N_{\bar{f} \circ \sigma_x} \circ \sigma_x^{-1}(x), \quad \sigma_x = \varphi \circ \phi_{\varphi^{-1}(x)} \circ \theta_{-x}, \quad \bar{f} = f \circ \varphi^{-1}. \quad (43)$$

Assume $N_{\bar{f}}$ converges with rate K to x^* . Then $\tilde{N}(x^*) = x^*$. The trick for seeing how \tilde{N} converges locally to x^* is to use (42) to remove the coordinate change σ_{x^*} from (43). Precisely, the iterative method N_g in (43) is replaced by the iterative method $\bar{N} = \sigma_{x^*} \circ N_{g \circ \sigma_{x^*}} \circ \sigma_{x^*}^{-1}$. By Lemma 24, this change will not alter the rate K of convergence provided $K > 1$. (Recall from Appendix A that the $K = 1$ case is more delicate.) Thus,

$$\tilde{N}(x) = \psi_x \circ \bar{N}_{\bar{f} \circ \psi_x} \circ \psi_x^{-1}(x), \quad \psi_x = \varphi \circ \phi_{\varphi^{-1}(x)} \circ \theta_{x^* - x} \circ \phi_{\varphi^{-1}(x^*)}^{-1} \circ \varphi^{-1}. \quad (44)$$

As arranged, ψ_{x^*} is the identity. If N , and hence \bar{N} , is reasonably nice then the radius of convergence — equivalently, the constant κ in Section 2 — associated with (44) should remain bounded if ψ_x remains sufficiently close to the identity. Indeed, all (44) is doing is applying \bar{N} to the cost function \bar{f} in the coordinate system determined by ψ_x . At the end of the day, lifting iterative methods to manifolds relies on this one simple principle: that the iterative method be robust to changes of coordinates.

9. Conclusion

The Newton method (1) is traditionally lifted to manifolds by endowing the manifold with a Riemannian structure and using (2). This strategy provides limited insight and may have a high computational cost when implemented. This motivates the study, from first principles, of lifting iterative methods from Euclidean space to manifolds.

Coordinate changes play a central role. Changing coordinates at each iteration is a novel yet easily understood and applied technique for enhancing the performance of iterative methods in Euclidean space (Section 4.3). Robustness to coordinate changes is key to lifting iterative methods to manifolds in useful ways (Section 8).

Newton methods on manifolds are defined customarily as iteration functions $E_f: M \rightarrow M$. This is unnecessarily restrictive; allowing E_f to depend on past history leads to path-dependent Newton methods (Section 6), and a change in focus from devising parametrisations to devising transformations (Section 7). The simplifications stemming from this generalisation are a consequence of eliminating the need for local lifts to agree globally; global agreement is a topological problem with little bearing on the computational problem of iteratively finding a critical point. While smooth global lifts of the Newton method always exist, global agreement may not be possible if the iterative method is not invariant to affine transformations and the manifold is non-parallelisable (Section 5.1).

The proposed framework for lifting the Newton method to manifolds is perhaps the most general one possible (Section 5): the condition on the cost function in Theorem 11 agrees with the necessary and sufficient condition in Theorem 1 for the Euclidean case, and it is difficult to see how conditions C1–C2 can be weakened (see Section 4.1).

Appendix A. Rate of Convergence of Iterates on Manifolds

Prior to this work¹, it was natural to define convergence with respect to a Riemannian metric. The belief that Newton methods should not depend on any Riemannian geometry led to the following. Compared with [1, Section 4.5], the lemmata here are careful to ensure the iterates do not fall outside the domain of definition of the iteration function.

Convergence rates are not preserved by arbitrary homeomorphisms. A sufficient condition for rates $K > 1$ is the following.

Lemma 26. *Let N be an iteration function on \mathbb{R}^n which converges locally to x^* with rate $K > 1$ and constant κ . Let U be a neighbourhood of x^* and $\phi: U \rightarrow V \subset \mathbb{R}^n$ a bi-Lipschitz homeomorphism about x^* , meaning there exist*

¹The main results of this paper were obtained in 2004–2005 and communicated privately to colleagues.

positive constants $\alpha, \beta \in \mathbb{R}$ such that

$$\forall x \in U, \quad \frac{1}{\alpha} \|x - x^*\| \leq \|\phi(x) - \phi(x^*)\| \leq \beta \|x - x^*\|. \quad (\text{A.1})$$

Then $\tilde{N} = \phi \circ N \circ \phi^{-1}$ converges locally to $\phi(x^*)$ with rate K and constant $\alpha^K \beta \kappa$.

Proof. As noted in Section 2, since N converges locally to x^* , for all sufficiently small balls B centred at x^* , N is defined on B , and $x \in B$ implies $N(x) \in B$ and $\|N(x) - x^*\| \leq \kappa \|x - x^*\|^K$. Choose such a B contained in U . Since ϕ is a homeomorphism, $Y = \phi(B)$ is a non-empty open subset of V . If $y \in Y$ then $\tilde{N}(y)$ is well-defined and contained in Y , and $\|\tilde{N}(y) - \phi(x^*)\| \leq \beta \|N(\phi^{-1}(y)) - x^*\| \leq \beta \kappa \|\phi^{-1}(y) - x^*\|^K \leq \alpha^K \beta \kappa \|y - \phi(x^*)\|^K$. \square

A significantly stronger condition is required if $K = 1$. One such example is the following.

Lemma 27. *Let N be an iteration function on \mathbb{R}^n converging locally to x^* at a linear rate. Let U be a neighbourhood of x^* and $\phi: U \rightarrow V \subset \mathbb{R}^n$ a \mathcal{C}^1 -diffeomorphism whose differential $D\phi$ at x^* is proportional to the identity. Then $\tilde{N} = \phi \circ N \circ \phi^{-1}$ converges locally to $\phi(x^*)$ at a linear rate.*

Proof. Let $\gamma \in \mathbb{R}$ be such that $D\phi(x^*) \cdot \xi = \gamma \xi$ for $\xi \in \mathbb{R}^n$. Note $\gamma \neq 0$ because ϕ is a diffeomorphism. Since $\phi(x) - \phi(x^*) = \gamma(x - x^*) + r(x)$ where $\lim_{x \rightarrow x^*} \|r(x)\|/\|x - x^*\| = 0$, by shrinking U to become a sufficiently small neighbourhood of x^* , it can be arranged for (A.1) to hold with $\beta = |\gamma| + \epsilon$ and $\alpha = \frac{1}{|\gamma| - \epsilon}$ for any $\epsilon > 0$. The result follows from Lemma 26 by choosing ϵ so that $\alpha \beta \kappa < 1$, where $\kappa < 1$ is the constant associated with N . \square

The above suggests the following definition. An iteration function $E: M \rightarrow M$ on an n -dimensional manifold M is said to converge locally with rate $K \geq 1$ to p^* with respect to the homeomorphism $\varphi: W \subset M \rightarrow V \subset \mathbb{R}^n$, where $p^* \in W$, if $\varphi \circ E \circ \varphi^{-1}$, as an iteration function on \mathbb{R}^n , converges locally with rate K to $\varphi(p^*)$.

If $K = 1$ or M is only a topological manifold, there is no distinguished choice of homeomorphism φ with respect to which convergence can be defined. If M is a \mathcal{C}^1 -manifold, $K > 1$ and an iterate converges with respect to one coordinate chart φ then Lemma 26 implies it converges with respect to any other chart ψ . (Proof: If $N = \varphi \circ E \circ \varphi^{-1}$ converges then, since $\psi \circ \varphi^{-1}$ is \mathcal{C}^1 and hence bi-Lipschitz on a possibly smaller domain, $\psi \circ E \circ \psi^{-1} = (\psi \circ \varphi^{-1}) \circ (\varphi \circ E \circ \varphi^{-1}) \circ (\psi \circ \varphi^{-1})^{-1}$ converges too.) Definition 28 affords a coordinate independent definition of rate of convergence.

Definition 28. An iteration function $E: M \rightarrow M$ on a \mathcal{C}^1 -differentiable manifold converges locally with rate $K > 1$ to $p^* \in M$ if there exists a coordinate chart $\varphi: W \rightarrow V \subset \mathbb{R}^n$ defined on a neighbourhood of p^* such that $\varphi \circ E \circ \varphi^{-1}$ converges locally with rate K to $\varphi(p^*)$ as an iteration function on \mathbb{R}^n .

Appendix B. Local Parametrisations

The normalisation $\phi_x(x) = x$ used in Section 4.1 does not generalise well to the manifold setting. Section 5 implicitly introduced $h_x(y) = \phi_x(x+y)$, thereby changing the normalisation to $h_x(0) = x$. Properties H1 to H3 of Section 5 are the analogues of properties P1 and P2 in Section 4.1.

Choosing $h_x(y) = x + y + y^2$ if x is rational and $h_x(y) = x + y - y^2$ if x is irrational exemplifies H1–H3 do not imply continuity of h . Conversely, h being \mathcal{C}^1 -smooth and satisfying H1 and H2 need not imply H3.

Example 29. Let $\alpha: \mathbb{R} \rightarrow \mathbb{R}$ be a \mathcal{C}^2 -smooth (or even \mathcal{C}^∞ -smooth) bump function satisfying: $0 \leq \alpha(t) \leq 1$; $\alpha(t) = \alpha'(t) = 0$ for $t \notin (1/2, 1)$; $\alpha(3/4) = 1$. Let $h(x, y) = x + y + x^{-1/2}\alpha(y/x^2)y^2$ if $x > 0$ and $h(x, y) = x + y$ otherwise. Then differentiation shows that $h(x, y)$ is \mathcal{C}^1 -smooth in (x, y) . Furthermore, $h_x(0) = x$, $Dh_x(0) = 1$ and $D^2h_x(0) = 0$. Therefore, H1 and H2 are satisfied, but H3 is not; if $x_n \rightarrow 0$ with $x_n > 0$ and $y_n = (3/4)x_n^2$ then $(h_{x_n}(y_n) - x_n - y_n)y_n^{-2} \rightarrow \infty$.

Nevertheless, a corollary of Lemma 30 is that h being \mathcal{C}^2 -smooth, or even just $D^2h_x(y)$ being continuous in (x, y) , suffices for H1 to imply H2 and H3.

Lemma 30. *If, for $x, y \in B(0; \rho)$, $D^2h_x(y)$ is bounded in (x, y) and continuous in y (that is, for each x , $h_x(y)$ is \mathcal{C}^2 -smooth in y) then h satisfying $H1_\rho$ implies it satisfies $H2_\rho$ and $H3_\rho$.*

Proof. Let $\alpha = \sup_{x, y \in B(0; \rho)} \|D^2h_x(y)\|$; then $H2_\rho$ is satisfied. Taylor's theorem implies $h_x(y) = x + y + \frac{1}{2}D^2h_x(x + t(y-x)) \cdot (y-x, y-x)$ for some $t \in [0, 1]$. Thus, $H3_\rho$ holds with $\beta = \alpha/2$. \square

If $h(y) = y + t^3 \sin(1/t)$ then $|h(y) - y| \leq |y|^2$ whenever $|y| \leq 1$, however, $D^2h(0)$ does not exist. This puts Lemma 31 into context.

Lemma 31. *If h satisfies $H3_\rho$ then it satisfies $H1_\rho$, and if additionally $D^2h_x(0)$ exists for $x \in B(0; \rho)$ then h satisfies $H2_\rho$ (with $\alpha = 2\beta$).*

Proof. That $H3_\rho$ implies $H1_\rho$ is clear. If $D^2h_x(0)$ exists, it is known that

$$\lim_{\|y\| \rightarrow 0} \|h_x(y) - 2h_x(0) + h_x(-y) - D^2h_x(0) \cdot (y, y)\| \|y\|^{-2} = 0. \quad (\text{B.1})$$

Thus, for any $\epsilon > 0$ there is a $\delta > 0$ such that $\|h_x(y) - 2x + h_x(-y) - D^2h_x(0) \cdot (y, y)\| \leq \epsilon \|y\|^2$ whenever $\|y\| \leq \delta$. Then $\|D^2h_x(0) \cdot (y, y)\| \leq \epsilon \|y\|^2 + \|h_x(y) - x - y\| + \|h_x(-y) - x - (-y)\| \leq (\epsilon + 2\beta)\|y\|^2$, proving the result; both sides scale as $\|y\|^2$ and $\epsilon > 0$ was arbitrary. \square

Lemma 32 asserts that $H3_\rho$ is preserved under second-order changes to h ; the straightforward proof is omitted.

Lemma 32. *For some $\rho > 0$, assume h satisfies $H3_\rho$. If there exists a $\gamma \in \mathbb{R}$ such that \tilde{h} satisfies $\|h_x(y) - \tilde{h}_x(y)\| \leq \gamma \|y\|^2$ whenever $x, y \in B(0; \rho)$ then \tilde{h} satisfies $H3_\rho$.*

The following two technical lemmata will be required in subsequent proofs; Lemma 33 is well-known.

Lemma 33. *Given $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\delta > 0$, define $L = \sup_{z \in B(0; \delta)} \|Dg(z)\|$ and $M = \sup_{z \in B(0; \delta)} \frac{1}{2} \|D^2g(z)\|$. If g is \mathcal{C}^1 -smooth on $B(0; \delta)$ and L is finite then $\|g(x) - g(y)\| \leq L\|x - y\|$, and if g is \mathcal{C}^2 -smooth on $B(0; \delta)$ and M is finite then $\|g(x) - g(y) - Dg(y) \cdot (x - y)\| \leq M\|x - y\|^2$ and $\|Dg(x) - Dg(y)\| \leq 2M\|x - y\|$ for $x, y \in B(0; \delta)$. If g is \mathcal{C}^1 -smooth on $\overline{B}(0; \delta)$, meaning it is \mathcal{C}^1 -smooth on an open set $U \supset \overline{B}(0; \delta)$, then L is finite, and M is finite if g is \mathcal{C}^2 -smooth on $\overline{B}(0; \delta)$.*

Lemma 34. *Fix a dimension n . Given scalars $\rho_1, \rho_2, \beta_1, G, L, M > 0$, there exist $\rho, \beta > 0$ such that, for any $\bar{h}: B_n(0; \rho_1) \rightarrow \mathbb{R}^n$ satisfying $\|\bar{h}(y) - y\| \leq \beta_1\|y\|^2$ for $y \in B(0; \rho_1)$, and for any $g: B_n(0; \rho_2) \rightarrow \mathbb{R}^n$ that is a \mathcal{C}^2 -diffeomorphism onto its image and satisfies $g(0) = 0$, $\|[Dg(0)]^{-1}\| \leq G$, $\|Dg(0)\| \leq L$ and $\sup_{z \in B(0; \rho_2)} \frac{1}{2} \|D^2g(z)\| \leq M$, it follows that $\tilde{h}(y) = (g \circ \bar{h} \circ [Dg(0)]^{-1})(y)$ is defined for $y \in B(0; \rho)$ and satisfies $\|\tilde{h}(y) - y\| \leq \beta\|y\|^2$.*

Proof. For brevity, define $A = [Dg(0)]^{-1}$. By successively shrinking $\rho > 0$ as required, the following requirements can be met for all $y \in B(0; \rho)$: $\|Ay\| \leq \rho G < \rho_1$; $\|\bar{h}(Ay) - Ay\| \leq \beta_1\|Ay\|^2 \leq \beta_1G^2\|y\|^2$; $\|A^{-1}\bar{h}(Ay) - y\| \leq \beta_1LG^2\|y\|^2$; $\|\bar{h}(Ay)\| \leq (1 + \beta_1\rho G)G\|y\| < \rho_2$; $\|g(\bar{h}(Ay)) - A^{-1}\bar{h}(Ay)\| \leq M\|\bar{h}(Ay)\|^2 \leq MG^2(1 + \rho\beta_1G)^2\|y\|^2$ (Lemma 33); and finally $\|\tilde{h}(y) - y\| \leq \|\bar{h}(y) - A^{-1}\bar{h}(Ay)\| + \|A^{-1}\bar{h}(Ay) - y\| \leq MG^2(1 + \rho\beta_1G)^2\|y\|^2 + \beta_1LG^2\|y\|^2$. Importantly, an appropriate value of ρ can be determined as a function of the other scalars and does not depend on g or \bar{h} . Similarly, $\beta = MG^2(1 + \rho\beta_1G)^2 + \beta_1LG^2$ suffices. \square

In certain situations, such as in Section 7.2, h_x is constructed from transformed versions of a prototype \bar{h} , as in Lemma 35.

Lemma 35. *Let $\bar{h}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ restricted to $B(0; \rho_1)$ satisfy $\|\bar{h}(y) - y\| \leq \beta\|y\|^2$ for some $\beta \in \mathbb{R}$. Assume $D^2\bar{h}(0)$ exists. Define $h_x(y) = g_x \circ \bar{h}([Dg_x(0)]^{-1} \cdot y)$ where, for each $x \in B(0; \rho_3) \subset \mathbb{R}^n$, $g_x: \mathbb{R}^n \rightarrow \mathbb{R}^n$ restricted to $B(0; \rho_2)$ is a \mathcal{C}^2 -diffeomorphism satisfying $g_x(0) = x$, $\|Dg_x(0)\| \leq L$ and $\|[Dg_x(0)]^{-1}\| \leq G$ for some $G, L \in \mathbb{R}$. Assume $M = \sup_{x \in B(0; \rho_3), y \in B(0; \rho_2)} \frac{1}{2} \|D^2g_x(y)\|^2 < \infty$. Here, $\rho_1, \rho_2, \rho_3 > 0$. Then h satisfies H1, H2 and H3.*

Proof. Since $D^2\bar{h}(0)$ exists and g_x is \mathcal{C}^2 -smooth, $D^2h_x(0)$ exists. By Lemma 31, it suffices to prove h satisfies H3. Fix $x \in B(0; \rho_3)$. Choose ρ and β as in Lemma 34 with $g(z) = g_x(z) - x$. Then $\|h_x(y) - x - y\| = \|(g \circ \bar{h} \circ [Dg(0)]^{-1})(y) - y\| \leq \beta\|y\|^2$ whenever $\|y\| < \rho$. Therefore h satisfies $H3_{\min\{\rho, \rho_2\}}$. \square

Properties H1–H3 are preserved under a change of coordinates.

Lemma 36. *Let $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ restricted to $B(0; \rho_3)$ be a \mathcal{C}^2 -diffeomorphism onto its image, with $g(0) = 0$. Given a function $h: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, define $\tilde{h}_x(y) = g \circ h_{g^{-1}(x)}(D(g^{-1})(x) \cdot y)$. If h satisfies H1 and H2 then \tilde{h} satisfies H1 and H2. If h satisfies H3 then \tilde{h} satisfies H3.*

Proof. Assume first that h satisfies $H3_{\rho_1}$. Choose a ρ_2 such that $0 < \rho_2 < \frac{\rho_3}{2}$ and $B(0; \rho_2) \subset g(B(0; \min\{\rho_1, \frac{\rho_3}{2}\}))$. Fix an $x \in B(0; \rho_2)$. Define $\tilde{h}(y) = h_{g^{-1}(x)}(y) - g^{-1}(x)$ and $\tilde{g}(z) = g(z + g^{-1}(x)) - x$. Note that $\tilde{h}(y)$ is well-defined for $\|y\| < \rho_1$ and $\tilde{g}(z)$ is well-defined for $\|z\| < \rho_2$. Then Lemma 34 is applicable, with \tilde{g} replacing g . (By shrinking ρ_3 if necessary, it can be assumed the derivatives of \tilde{g} are uniformly bounded.) In particular, there exist ρ and β , independent of x , such that $\|\tilde{h}_x(y) - x - y\| = \|(\tilde{g} \circ \tilde{h} \circ [Dg(0)]^{-1})(y) - y\| \leq \beta \|y\|^2$ whenever $y \in B(0; \rho)$. Therefore \tilde{h} satisfies $H3_{\min\{\rho, \rho_2\}}$, as required.

Next, assume h satisfies H1 and H2 (but not necessarily H3). It is reasonably clear that $\tilde{h}_x(y)$ has a sufficiently large domain of definition required for $\tilde{h}_x(0)$, $D\tilde{h}_x(0)$ and $D^2\tilde{h}_x(0)$ to exist in a neighbourhood of $x = 0$. Explicit calculations, using the chain rule to compute derivatives, verify that \tilde{h} satisfies H1 and H2. \square

It is remarked that the tedious nature of the last few proofs comes from the necessity of ensuring the transformed h has a valid domain of definition. This is a consequence of the standing assumption that h itself need not be defined on the whole of $\mathbb{R}^n \times \mathbb{R}^n$. This becomes important when coordinate charts on manifolds enter the picture.

Appendix C. Further Results on the Generalised Newton Method

Appendix C.1. Intrinsic Conditions

Condition (37) does not depend on the choice of coordinates.

Proposition 37. *In Theorem 11, if (37) holds, it holds with respect to any \mathcal{C}^2 -chart $(\tilde{U}, \tilde{\varphi})$ with $\tilde{\varphi}(p^*) = 0$ and \tilde{U} sufficiently small.*

Proof. Referring to Theorem 11, let $(\tilde{U}, \tilde{\varphi})$ be a chart with $\tilde{\varphi}(p^*) = 0$ and choose $\rho > 0$ so that $h = \varphi \circ \tilde{\varphi}^{-1}$ is well-defined on $\overline{B}(0; \rho)$. Then $H_{f \circ \tilde{\varphi}^{-1}}(x) = H_{\tilde{f} \circ h}(x) = A_x^T H_{\tilde{f}}(h(x)) A_x + G_x$ where A_x and G_x are the matrix representations of Dh and $(Df \circ h)D^2h$ respectively. Since Dh and $Df \circ h$ are \mathcal{C}^1 -smooth and D^2h is continuous, there exist constants α, β such that $\|A_x - A_0\| \leq \alpha \|x\|$ and $\|G_x\| \leq \beta \|x\|$ whenever $x \in \overline{B}(0; \rho)$. Similarly, from (37) and Taylor series arguments, there exists a constant γ such that $\|[H_{\tilde{f}}(h(x)) - H_{\tilde{f}}(0)]A_0x\| \leq \|[H_{\tilde{f}}(h(x)) - H_{\tilde{f}}(0)]h(x)\| + \|[H_{\tilde{f}}(h(x)) - H_{\tilde{f}}(0)](h(x) - A_0x)\| \leq \gamma \|x\|^2$ whenever $x \in \overline{B}(0; \rho)$. Shrink \tilde{U} to equal $\tilde{\varphi}^{-1}(B(0; \rho))$. The result follows by noting

$$\begin{aligned} \|[H_{f \circ \tilde{\varphi}^{-1}}(x) - H_{f \circ \tilde{\varphi}^{-1}}(0)]x\| &\leq \|[A_x^T H_{\tilde{f}}(h(x)) A_x - A_x^T H_{\tilde{f}}(h(x)) A_0]x\| + \\ &\|[A_x^T H_{\tilde{f}}(h(x)) A_0 - A_0^T H_{\tilde{f}}(h(x)) A_0]x\| + \|[A_0^T H_{\tilde{f}}(h(x)) A_0 - A_0^T H_{\tilde{f}}(0) A_0]x\| + \\ &\|G_x x\|. \end{aligned} \quad (\text{C.1})$$

\square

Conditions C1–C2 are also intrinsic; the choice of coordinate charts is immaterial and the conditions are preserved under diffeomorphisms. Let $h_*: TM \rightarrow TN$ denote the push-forward of tangent vectors induced by a map $h: M \rightarrow N$ between manifolds; $h_*(v_p) = Dh(p) \cdot v_p$.

Proposition 38. *Let $\phi, \psi: TM \rightarrow M$ satisfy C1–C2. Then about any point $p \in M$, C1 and C2 hold with respect to any \mathcal{C}^2 -chart $(\tilde{U}, \tilde{\varphi})$ with $\tilde{\varphi}(p) = 0$. Furthermore, if $h: M \rightarrow N$ is a \mathcal{C}^2 -diffeomorphism of manifolds then the induced maps $\tilde{\phi} = h \circ \phi \circ h_*^{-1}$ and $\tilde{\psi} = h \circ \psi \circ h_*^{-1}$ satisfy C1–C2.*

Proof. Let $h: M \rightarrow N$ be a \mathcal{C}^2 -diffeomorphism. Fix $p \in M$. Let $(\tilde{U}, \tilde{\varphi})$ be a \mathcal{C}^2 -chart on N with $\tilde{\varphi} \circ h(p) = 0$. It will be shown $\hat{\phi} = \tilde{\varphi} \circ \tilde{\phi} \circ \tau_{\tilde{\varphi}}^{-1}$ satisfies H1 and H2, and $\hat{\psi} = \tilde{\varphi} \circ \tilde{\psi} \circ \tau_{\tilde{\varphi}}^{-1}$ satisfies H3. This proves the second part of the lemma. The first part then follows by letting $h: M \rightarrow M$ be the identity map.

Let (U, φ) be a \mathcal{C}^2 -chart on M with $\varphi(p) = 0$ and such that $\hat{\phi} = \varphi \circ \phi \circ \tau_{\varphi}^{-1}$ satisfies H1 and H2, and $\hat{\psi} = \varphi \circ \psi \circ \tau_{\varphi}^{-1}$ satisfies H3. Let $g = \tilde{\varphi} \circ h \circ \varphi^{-1}$; it is a \mathcal{C}^2 -diffeomorphism from $\varphi(U \cap h^{-1}(\tilde{U}))$ to $\tilde{\varphi}(h(U) \cap \tilde{U})$ and $\hat{\psi}(x, y) = g \circ \hat{\psi}_{g^{-1}(x)} \circ D(g^{-1})(x) \cdot y$. Apply Lemma 36 to conclude $\hat{\psi}$ satisfies H3. Analogously, Lemma 36 implies $\hat{\phi}$ satisfies H1 and H2. \square

Appendix C.2. Sufficient Conditions

Conditions C1–C2 are readily satisfied by \mathcal{C}^2 -smooth parametrisations. In this case M must be \mathcal{C}^3 -smooth. If M were only \mathcal{C}^2 -smooth then $\phi: TM \rightarrow M$ at best can be \mathcal{C}^1 -smooth because TM is only a \mathcal{C}^1 -manifold.

Lemma 39. *Let M be a \mathcal{C}^3 -manifold. If ϕ is \mathcal{C}^2 -smooth and, for all $p \in M$, $\phi_p(0_p) = p$ and $D\phi_p(0_p) = I$, then C1 holds. If ψ is \mathcal{C}^2 -smooth and, for all $p \in M$, $\psi_p(0_p) = p$ and $D\psi_p(0_p) = I$, then C2 holds.*

Proof. Follows from Lemma 30. \square

Remark 40. Since C1–C2 are local in nature (Lemma 42), it suffices in Lemma 39 for ϕ and ψ to be smooth on a neighbourhood of the zero section of TM .

Conditions C1–C2 are preserved under restriction to submanifolds.

Lemma 41. *Let $i: N \rightarrow M$ be a \mathcal{C}^2 -embedding of N in M , with $i_*: TN \rightarrow TM$ the induced push-forward of tangent vectors. Let $\phi, \psi: TM \rightarrow M$ be parametrisations of M satisfying C1–C2, and $\tilde{\phi}, \tilde{\psi}: TN \rightarrow N$ parametrisations of N satisfying $\phi \circ i_* = i \circ \tilde{\phi}$ and $\psi \circ i_* = i \circ \tilde{\psi}$. Then $\tilde{\phi}, \tilde{\psi}$ satisfy C1–C2.*

Proof. From Proposition 38, it suffices to assume $N \subset M$. Then $\tilde{\phi}$ and $\tilde{\psi}$ are simply the restrictions of ϕ and ψ to TN . The result follows by observing that if $h: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ in the definitions of H1–H3 is restricted to $\mathbb{R}^m \subset \mathbb{R}^n$ then H1–H3 would continue to hold. \square

One way to express precisely the local nature of C1–C2 is with the aid of a Riemannian metric on M .

Lemma 42. *Let $\phi, \psi: TM \rightarrow M$ satisfy C1–C2 where M is a \mathcal{C}^2 -Riemannian manifold. Let $r: M \rightarrow (0, \infty)$ be a possibly discontinuous function. Assume $\tilde{\phi}, \tilde{\psi}: TM \rightarrow M$ satisfy $\tilde{\phi}(v_p) = \phi(v_p)$ and $\tilde{\psi}(v_p) = \psi(v_p)$ whenever $\|v_p\| < r(p)$. Then $\tilde{\phi}$ satisfies C1. If $\inf_{p \in K} r(p) > 0$ for any compact $K \subset M$ then $\tilde{\psi}$ satisfies C2.*

Proof. Fix $p \in M$ and let φ and ρ be such that C1 and C2 hold. Then $B(0; \rho) \times B(0; \rho)$ is in the image of τ_φ ; let V be its pre-image. For any $\bar{r} > 0$, the set $V_{\bar{r}} = \{v_p \in V \mid \|v_p\| < \bar{r}\}$ is open, hence $\tau_\varphi(V_{\bar{r}})$ is open too.

Choose an $x \in B(0; \delta)$ and let $\bar{r} = r(\varphi^{-1}(x))$. There exists a $\delta_x > 0$ such that $(x, B(0; \delta_x)) \subset \tau_\varphi(V_{\bar{r}})$. Restricted to $(x, B(0; \delta_x))$, $\varphi \circ \phi \circ \tau_\varphi^{-1}$ and $\varphi \circ \tilde{\phi} \circ \tau_\varphi^{-1}$ are equal. It follows that ϕ satisfies C1.

Let $K = \varphi^{-1}(\bar{B}(0; \rho/2))$ and $\bar{r} = \inf_{p \in K} r(p)$. Let $\bar{\rho} \in (0, \rho/2)$ be such that $B(0; \bar{\rho}) \times B(0; \bar{\rho}) \subset \tau_\varphi(V_{\bar{r}})$. Restricted to $B(0; \bar{\rho}) \times B(0; \bar{\rho})$, $\varphi \circ \psi \circ \tau_\varphi^{-1}$ and $\varphi \circ \tilde{\psi} \circ \tau_\varphi^{-1}$ are equal. It follows that ψ satisfies C2. \square

Appendix C.3. Embedded Submanifolds of Euclidean Space

For manifolds embedded in Euclidean space, C1–C2 can be expressed in extrinsic coordinates.

Treating \mathbb{R}^m as a manifold, a parametrisation $\phi: T\mathbb{R}^m \rightarrow \mathbb{R}^m$ can be specified by its representation $\hat{\phi}: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ with respect to the identity chart, denoted $\phi = \hat{\phi} \circ \tau_I$. Given a \mathcal{C}^2 -embedding $i: M \rightarrow \mathbb{R}^m$, let $V_x M$ for $x \in i(M)$ denote the realisation of $T_{i^{-1}(x)}M$ as a subspace of \mathbb{R}^m , that is, $(x, V_x M) = \tau_I \circ i_*(T_{i^{-1}(x)}M)$ where $i_*: TM \rightarrow T\mathbb{R}^m$ is the push-forward of i . (The elements of $V_x M$ are the vectors $\gamma'(0)$ where $\gamma: (-\epsilon, \epsilon) \rightarrow \mathbb{R}^m$, $\gamma(0) = x$, is a curve whose image is contained in $i(M)$.)

If $\hat{\phi}(x, y)$ belongs to $i(M)$ whenever $x \in i(M)$ and $y \in V_x M$ then it induces a parametrisation $\tilde{\phi}: TM \rightarrow M$ given by $\tilde{\phi} = i^{-1} \circ \hat{\phi} \circ \tau_I \circ i_*$. In essence, $\hat{\phi}$ maps a point $x + y$ on the affine tangent space of $i(M)$ at x , to the point $\tilde{\phi}(x, y)$ on $i(M)$. This is how parametrisations were specified in [15].

Lemma 43. *Let $i: M \rightarrow \mathbb{R}^m$ be a \mathcal{C}^2 -embedding of a manifold M . With notation as above, assume $\hat{\phi}, \hat{\psi}: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ satisfy: $\forall z \in i(M), \exists \alpha, \beta, \rho \in \mathbb{R}$ with $\rho > 0, \forall x \in B(z; \rho) \cap i(M), \forall y \in B(0; \rho) \cap V_x M, \hat{\phi}_x(y), \hat{\psi}_x(y) \in i(M), \hat{\phi}_x(0) = x, D\hat{\phi}_x(0) \cdot y = y, \|D^2\hat{\phi}_x(0) \cdot (y, y)\| \leq \alpha \|y\|^2, \|\hat{\psi}_x(y) - x - y\| \leq \beta \|y\|^2$. Then the parametrisations $\tilde{\phi}, \tilde{\psi}: TM \rightarrow M$ defined by $\tilde{\phi} = i^{-1} \circ \hat{\phi} \circ \tau_I \circ i_*$ and $\tilde{\psi} = i^{-1} \circ \hat{\psi} \circ \tau_I \circ i_*$ satisfy C1 and C2 of Section 5.*

Proof. For $x \in i(M)$, let $P_x: \mathbb{R}^m \rightarrow V_x M$ denote Euclidean projection onto $V_x M$. Extend $\hat{\phi}$ by defining $\hat{\phi}(x, y) = x + y$ for $x \notin i(M)$, and $\hat{\phi}(x, y) = \hat{\phi}(x, P_x(y)) + y - P_x(y)$ for $x \in i(M)$ and $y \notin V_x M$. Extend $\hat{\psi}$ similarly. Then $\phi = \hat{\phi} \circ \tau_I$ and $\psi = \hat{\psi} \circ \tau_I$ satisfy C1–C2. (Fix $p \in \mathbb{R}^m$ and define $\varphi(x) = x - p$. Note $\tau_I \circ \tau_\varphi^{-1}(x, y) = (x + p, y)$. Hence $\varphi \circ \phi \circ \tau_\varphi^{-1}(x, y) = \hat{\phi}(x + p, y) - p$. Same for ψ . It is readily verified the assumptions in the proposition ensure H1, H2 and H3 are satisfied.) Hence, from Lemma 41, $\tilde{\phi}$ and $\tilde{\psi}$ satisfy C1–C2. \square

Acknowledgements

This work was funded in part by the Australian Research Council. Special thanks to Dr Jochen Trumpf for insightful and thought-provoking discussions during the preliminary stages of this paper, and to the two anonymous reviewers for excellent guidance on improving the presentation.

References

- [1] P. A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.
- [2] P. A. Absil, R. Mahony, R. Sepulchre, and P. Van Dooren. A Grassmann-Rayleigh quotient iteration for computing invariant subspaces. *SIAM Review. A Publication of the Society for Industrial and Applied Mathematics*, 44(1):57–73, 2002.
- [3] P. A. Absil and J. Malick. Projection-Like Retractions on Matrix Manifolds. *Siam Journal on Optimization*, 22(1):135–158, 2012.
- [4] P. A. Absil, R. Sepulchre, P. Van Dooren, and R. Mahony. Cubically convergent iterations for invariant subspace computation. *SIAM Journal on Matrix Analysis and Applications*, 26(1):70–96, 2004.
- [5] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub. Newton’s method on Riemannian manifolds and a geometric model for the human spine. *Ima Journal of Numerical Analysis*, 22(3):359–390, 2002.
- [6] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton’s method in Riemannian manifolds. *Foundations of Computational Mathematics*, 8(2):197–226, 2008.
- [7] I. K. Argyros. An improved unifying convergence analysis of Newton’s method in Riemannian manifolds. *Journal of Applied Mathematics and Computing*, 25(1-2):345–351, 2007.
- [8] P. Deuffhard. *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*. Springer Series in Computational Mathematics. Springer, 2004.
- [9] A. Edelman, T. A. Arias, and S. T. Smith. The Geometry of Algorithms with Orthogonality Constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, Jan. 1998.
- [10] O. Ferreira and B. Svaiter. Kantorovich’s theorem on Newton’s method in Riemannian Manifolds. *Journal of Complexity*, 18(1):304–329, 2002.
- [11] D. Gabay. Minimizing a differentiable function over a differentiable manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982.
- [12] U. Helmke and J. B. Moore. *Optimization and dynamical systems*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1994.
- [13] M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, 1974.

- [14] L. Kantorovich and G. Akhilov. *Functional Analysis in Normed Spaces*. Fizmatgiz, Moscow, 1959.
- [15] J. H. Manton. Optimisation algorithms exploiting unitary constraints. *IEEE Transactions on Signal Processing*, 50(3):635–650, March 2002.
- [16] J. H. Manton. Optimisation geometry. In K. Hüper and J. Trumpf, editors, *Mathematical System Theory — Festschrift in Honor of Uwe Helmke on the Occasion of his Sixtieth Birthday*, pages 261–274. CreateSpace, 2013.
- [17] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, 1970.
- [18] E. Polak. *Optimization: Algorithms and Consistent Approximations*. Springer-Verlag, 1997.
- [19] M. Shub. Some remarks on dynamical systems and numerical analysis. In *Dynamical systems and partial differential equations (Caracas, 1984)*, pages 69–91. Univ. Simon Bolivar, Caracas, 1986.