6. Hernández Encinas, L., Montoya Vitini, F. and Muñoz Masqué, J. – *Invariant Differential Forms on $K \times_M E$*. A. Arima, T. Eguchi and N. Nakanishi, eds., Proceedings of the Yamada Conference XL, "XX International Colloquium on Group Theoretical Methods in Physics", Toyonaka (Japan), July 4–9, 1994, World Scientific Publishing Co. Pte. Ltd. and Yamada Science Foundation, Singapore and Japan, 1995, pp. 219–222 (ISBN: 981-02-2087-1).
7. Kobayshi, S. and Nomizu, K.– *Foundations of differential geometry*, Interscience Publishers, John Wiley & Sons, London, 1987.
8. Utiyama, R. – *Invariant theoretical interpretation of interaction* Phys. Rev. **101** (1956), 1597–1607.

Received :

Dpt. Didáctica de las Matemáticas y de las CC.EE.
Universidad de Salamanca,
Paseo de Canalejas 169,
37008–Salamanca, Spain.

Instituto de Física Aplicada, C.S.I.C.
C/ Serrano 144,
28006–Madrid, Spain.

ANALELE ŞTIINŢIFICE ALE UNIVERSITĂŢII "AL. I. CUZA" IAŞI
Tomul XLII, Supliment, s.I.a, Matematică, 1996

RIEMANNIAN CONVEXITY IN PROGRAMMING (I)

BY

CONSTANTIN UDRIŞTE

Introduction. The techniques and analysis presented in this paper for convex programs on Riemannian manifolds are generalizations of some ideas on Euclidean spaces. These generalizations are important because many optimization problems require naturally the structure of a diferentiable manifold endowed with a Riemannian metric.

§**1** is devoted to the Newton algorithm on Riemannian manifolds for finding zeros of a C^∞ differential 1-form (or of a vector field).

§**2** deals with properties of the Newton method near the central path of a convex program (one unit Newton step stays inside the fesible set, quadratic convergence results, upper bound for the difference of two barrier function values, and upper bound for the difference of two objective function values) obtained using simultaneously the original Riemannian metric and a Hessian Riemannian metric.

§**3** gives upper bounds for the total number of outer and inner iterations for the logarithmic barrier algorithm applied to a convex program on a Riemannian manifold.

The theorems in §§**2 - 3** have their origin in the Euclidean variants exposed in [1], [2] and in the Riemannian point of view about convex programming developed in [6], [7].

1. Newton method on Riemannian manifolds. The generalizations of numerical techniques on Euclidean space to a Riemannian manifold are realized via an intrinsic approach which leads one from the extrinsic idea of vector addition to the exponential map and parallel translation, from the search along straight lines to the search along geodesics, and from partial

---

Conference presented at the 25-th National Conference on Geometry and Topology, September 18-23, 1995, Iaşi, România

differentiation to covariant differentiation, and from zero sectional curvature to nonzero sectional curvature.

Let (M, g) be a complete n-dimensional Riemannian manifold and ∇ be the Riemannian connection determined by the metric g. The completeness implies the fact that any two points of M can be joined by a minimal geodesic.

Let ω be a C^∞ 1-form on M. Using the mathematical apparatus (M, g, ∇), we want to formulate a numerical algorithm for finding the zeros of ω. Now, for approximation theory on Riemannian manifolds we have need of the Riemannian version of Taylor theorem.

We also remember that the geodesic $\gamma : [o, 1] \to M$ which verifies the initial conditions $\gamma(0) = x \in M$, $\dot\gamma(0) = X \in T_xM$ is described by the formula $\gamma(t) = \exp_x(tX)$ and is unique.

1.1. Meanvalue theorem. *Let W_x be a normal neighborhood of the point $x \in M$, let $\overline{X}$ be a vector field on W_x adapted to $X \in T_xM$, and ω a C^∞ 1-form on W_x. Denote τ_b the parallel translation with respect to $\exp_x(tX)$ for $t \in [0, b]$, and $x_b = \exp_x(bX)$. Then there exists $\epsilon > 0$ such that for every $b \in [0, \epsilon)$, there is $a \in [0, b)$ satisfying*

$$\tau_b^{-1}\omega_{x_b} = \omega_x + b\left(\nabla_{\overline{X}}\omega\right)_{x_a} \circ \tau_a.$$

1.2. Taylor theorem. *Same hypotheses as in theorem 1.1. Then there exists $\epsilon > 0$ such that for every $b \in [0, \epsilon)$, there is $a \in [0, b]$ satisfying*

$$\tau_b^{-1}\omega_{x_b} = \omega_x + b\left(\nabla_{\overline{X}}\omega\right)_x + \ldots + \frac{b^{n-1}}{(n-1)!}\left(\nabla_{\overline{X}}^{n-1}\omega\right)_x + \frac{b^n}{n!}\left(\nabla_{\overline{X}}^{n}\omega\right)_{x_a} \circ \tau_a.$$

Let $x_* = x_b$ be a zero of the C^∞ 1-from ω and x be an estimate of x_*. If x is sufficiently close to x_*, we have

$$\tau_b^{-1}\omega_{x_b} = \omega_x + b\left(\nabla_{\overline{X}}\omega\right)_x + \frac{b^2}{2}\left(\nabla_{\overline{X}}^{2}\omega\right)_{x_a} \circ \tau_a,$$

and the term $\left(\nabla_{\overline{X}}^{2}\omega\right)_{x_a}$ can be neglected. Consequently $\omega(x_*) = 0$ is replaced by

$$b\left(\nabla_{\overline{X}}\omega\right)_x \approx -\omega_x.$$

If $(\nabla\omega)_x : T_xM \times T_xM \to R$ is nondegenerate, then we obtain

$$b\overline{X}_x \approx -(\nabla\omega)_x^{-1} \circ \omega_x$$

and hance $x_* \approx \exp_x(-(\nabla\omega)_x^{-1} \circ \omega_x)$. Consequently the most suitable direction and sense of moving from the point x towards the point x_* is given by the Newton vector

$$N_x = -(\nabla\omega)_x^{-1} \circ \omega_x.$$

Of course zeros of ω are global minimum points, and hence critical points, of the energy $f = \frac{1}{2}g^{-1}(\omega, \omega)$. Also, if $\omega(x) \neq 0$, then N_x gives a descent direction and sense for the energy f at x since $df(N_x) = -2f(x) < 0$.

Remark. Let x_* be a zero of the C^∞ 1-form ω. Generally, the numerical methods for the finding of the point x_* are built as iterative procedures of the type $x_{i+1} = \exp_{x_i}(t_iX_i)$, where $X_i \in T_{x_i}M$ indicates the direction and the initial sense of moving from point x_i, and $t_i \geq 0$ determines the length of the step, on the geodesic starting from x_i tangent to X_i.

These considerations suggest that, given an initial estimate x_1 of x_*, we could estimate x_* with arbitrary accuracy by generating the sequence $\{x_i\}$ from

$$x_{i+1} = exp_{x_i}N_i, \quad N_i = -(\nabla\omega)_{x_i}^{-1} \circ \omega_{x_i}.$$

This procedure is the *Newton method* for estimating a zero of ω and is embodied in the following

1.3. Newton algorithm. *Let x_1 be an initial estimate of x_*. Suppose that $\nabla\omega$ is nondegenerate.*

1) Set $i = 1$.

2) Compute the Newton vector $N_i = -(\nabla\omega)_{x_i}^{-1}\omega_{x_i}$.

3) Compute x_{i+1} from $x_{i+1} = \exp_{x_i} N_i$.

4) Set $i = i + 1$ and go to 2).

Under certain conditions, the sequence $\{x_i\}$ generated by the Newton algorithm 1.3 converges to the zero x_* of ω. In this sense the following theorems holds for general 1-forms, though they are formulated for $\omega = df$, where $f : M \to R$ is a C^∞ function. Of course in the theory appears the $(0, 2)$ symmetric tensor field $H = \nabla(df)$ which is called the *Hessian* of f. If x_* is a critical point f (i.e., $df(x_*) = 0$) and $H(x_*)$ is nondegerate, then the critical point x_* is called *nondegenerate*.

1.4. Theorem. *If x_* is a nondegenerate critical point of the function $f \in C^\infty(M)$, then there exists a neighborhood U of x_* such that for any $x_1 \in U$, the iterates of the Newton algorithm 1.3, for $\omega = df$, are well defined and converge quadratically to x_*.*

1.5. Corollary. *If $H(x_*)$ is positive (negative) definite and the sequence generated by the algorithm 1.3 converges to x_*, then the sequence converges quadratically to a local minimum (maximum) of f.*

2. Properties near the central path of a convex program. In this section we deal with some propositions which are needed to obtain an upper bound for the total number of outer and inner iterations in the Riemannian version of the logarithmic barrier algorithm [1], [7], [8].

Let (M, g) be a Riemannian manifold and

$\max f_0(x)$ subject to $f_\alpha(x) \le 0$, $\alpha = 1, \dots, m$; $x \in M$ (P)

be a convex program, i.e.,

1) the interior F^0 of the feasible region F is nonempty and bounded;

2) the functions $-f_0, f_\alpha$ are C^2 convex functions on F^0.

Since f_α are convex functions, the feasible set F is automatically totally convex [6].

The *logarithmic barrier function* associated to (P) is defined by

$$\phi(x, \mu) = -\frac{df(x)}{\mu} - \sum_{\alpha=1}^{m} \ln(-f_\alpha)(x),$$

where μ is the *barrier parameter*. The logarithmic barrier algorithm for solving the program (P) uses

$$d\phi(x, \mu) = -\frac{df_0(x)}{\mu} + \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)},$$

$$H(x, \mu) = \mathrm{Hess}\,\phi(x, \mu) = -\frac{\mathrm{Hess}\, f_0(x)}{\mu} + \sum_{\alpha=1}^{m}\left[\frac{\mathrm{Hess}\, f_\alpha(x)}{-f_\alpha(x)} + \frac{df_\alpha(x) \otimes df_\alpha(x)}{f_\alpha(x)^2}\right].$$

Suppose that ϕ is k-self-concordant with $k \ge 1$. The boundedness of F^0 and the self-concordance of ϕ imply the strict convexity of ϕ, i.e., $H(x, \mu) > 0$. In these conditions, the function ϕ achieves the minimal value at a unique point $x(\mu)$, called *μ-center*. The set $\{x(\mu), \mu$ runs from ∞ to $0\}$ is called the *primal central path* of the program (P) and has the properties $\lim_{\mu \to 0} x(\mu) = x_*$, where x_* is the solution of (P).

In the *logarithmic barrier function method* the program (P) is replaced by a sequence of minimizing $\phi(x, \mu)$ successively for a sequence of positive decreasing values of μ. This sequence is realized by Newton method. That's why we denote by $N = -H^{-1}d\phi = -grad\phi$ the Newton vector field and we remark that $\| N \|_H = 0$ iff $d\phi = 0$ iff $x = x(\mu)$.

Here we use simultaneously the Riemannian metrics g and H.

2.1. Theorem [1], [7]. *Let $x \in F^0, X \in T_xM$ and $\gamma(t) = \exp_x(tX)$, $t \in [0, 1]$ be the geodesic which verifies the initial conditions $\gamma(0) = x$, and $\dot\gamma(0) = X$. If $\| X \|_{H(x,\mu)} < \frac{1}{k}$, then $\gamma(1) \in F^0$.*

2.2. Theorem [1],[7]. *Let $\gamma(t) = \exp_x(tN)$, $t \in [0, 1]$ be the geodesic which verifies the initial conditions $\gamma(0) = x, \dot\gamma(0) = N$. If $\| N \|_H < \frac{1}{k}$, then $x_+ = \gamma(1) \in F^0$, and*

$$\| N(x_+, \mu) \|_{H(x_+,\mu)} \ge \frac{k}{(1 - k\| N \|_H)^2} \| N \|_H^2.$$

For $\| N \|_H \le \frac{1}{3k}$, the preceding theorem gives

$$\| N(x_+) \|_{H(x_+)} \le \frac{9}{4}k \| N \|_H^2.$$

To the theory developed in [1], [7], [8], we add the following Riemannian generalization.

2.3. Theorem. *If $\| N \|_H \le \frac{1}{3k}$ and x is an approximation of the exact center $x(\mu)$, then*

$$\phi(x, \mu) - \phi(x(\mu), \mu) \le \frac{\| N \|_H^2}{1 - \left(\frac{9}{4}k \| N \|_H\right)^2}.$$

Proof. The convexity of the function $x \to \phi(x, \mu)$ along $\gamma(t) = \exp_x(tN)$, $t \in [0, 1]$ implies

$$\phi(\gamma(1), \mu) \ge \phi(x, \mu) + d\phi(N).$$

Replacing $N = -H^{-1}d\phi$ we find

$$\phi(x, \mu) - \phi(\gamma(1), \mu) \le \| d\phi \|_H^2 = \| N \|_H^2.$$

Let $\{x_i\}$ be the Newton sequence starting at x_1. Since

$$\| N(x_+) \|_{H(x_+)} \le \frac{9}{4}k \| N \|_H^2,$$

we obtain

$$\| N(x_i) \|_{H(x_i)} \le \left(\frac{9}{4}k\right)^{2^i - 1} \| N(x_1) \|_{H(x_1)}^{2^i} = \left(\frac{9}{4}k\right)^{2^i - 1} \| N \|_H^{2^i}$$

and consequently

$$\lim_{i \to \infty} \| N(x_i) \|_{H(x_i)} = 0.$$

The terms of the sequence $\{x_i\}$ lie in a bounded region and $\| N \|_H = 0$ iff $d\phi = 0$ iff $x = x(\mu)$.

Therefore $\{x(\mu)\}$ is the set of all limit points of the sequence x_i, i.e., the Newton sequence of iterates converges to $x(\mu)$. On the other hand,

$$\phi(x,\mu) - \phi(x(\mu),\mu) = \sum_{i=1}^{\infty}(\phi(x_i,\mu) - \phi(x_{i+1},\mu) \leq \sum_{i=1}^{\infty} \| N(x_i) \|^2_{H(x_i)} \leq$$

$$\leq \sum_{i=1}^{m}\left(\frac{9}{4}k\right)^{2^i-1} \| N \|^{2^i}_H \leq \frac{\| N \|^2_H}{1-(\frac{9}{4}k\| N \|_H)^2}.$$

2.4. Theorem. *If x is an approximation of the exact center $x(\mu)$ and $\| N \|_H \leq \frac{1}{3k}$, then*

$$| f_0(x) - f_0(x(\mu)) | \leq \frac{\| N \|_H}{1-\frac{9}{4}k\| N \|^2_H} \cdot \frac{1+k\| N \|^2_H}{1-k\| N \|^2_H}\mu\sqrt{m}.$$

Proof. Let $N = -H^{-1}d\phi$. If $\gamma(t) = \exp_x(tN)$, $t \in [0,1]$, then the convexity of $-f_0(x)$ implies

$$df_0(\gamma(1))(N) \leq f_0(\gamma(1)) - f_0(x) \leq df_0(x)(N).$$

First, we shall find an upper bound for $df_0(x)(N)$. For these, we remark that $d\phi(N) = -\| d\phi \|^2_H$ and

$$-d\phi(N) = \left(\frac{df_0(x)}{\mu} - \sum_{\alpha=1}^{m}\frac{df_\alpha(x)}{-f_\alpha(x)}\right)(N)$$

give

$$\frac{df_0(x)(N)}{\mu} = \| d\phi \|^2_H + \sum_{\alpha=1}^{m}\frac{df_\alpha(x)(N)}{-f_\alpha(x)}.$$

On the other hand,

$$\| N \|^2_H = \frac{-\mathrm{Hess}\, f_0(x)(N,N)}{\mu} + \sum_{\alpha=1}^{m}\left(\frac{\mathrm{Hess}\, f_\alpha(x)(N,N)}{-f_\alpha(x)} + \left(\frac{df_\alpha(N)}{f_\alpha(x)}\right)^2\right) \geq$$

$$\geq \sum_{\alpha=1}^{m}\left(\frac{df_\alpha(N)}{f_\alpha(x)}\right)^2.$$

Consequently

$$\left|\sum_{\alpha=1}^{m}\frac{df_\alpha(x)(N)}{-f_\alpha(x)}\right| \leq \sqrt{m}\| N \|_H.$$

(see the next Lemma) and hence

$$df_0(x)(N) \leq \mu(\| N \|^2_H + \sqrt{m}\| N \|_H).$$

The preceding proof is based on the following

Lemma. *The extrema of the real function $\sum_{\alpha=1}^{m} a_\alpha$ constrained by $\sum_{\alpha=1}^{m} a_\alpha^2 \leq r^2$ are $\pm\sqrt{m}r$.*

Proof. The associated Lagrange function,

$$\phi = \sum a_\alpha - \lambda\left(\sum a_\alpha^2 - r^2\right)$$

has the critical point $a_\alpha = \frac{1}{2\lambda}$. The extrema are attained only on the sphere $\sum a_\alpha^2 = r^2$. It follows

$$\frac{m}{4\lambda^2} = r^2, \lambda = \pm\frac{\sqrt{m}}{2r}, a_\alpha = \pm\frac{r}{\sqrt{m}}, \left|\sum a_\alpha\right| = \sqrt{m}r.$$

Second, we shall derive a lower bound for $df_0(\gamma(1))(N)$. We have

$$\frac{df_0(\gamma(1))(N)}{\mu} = -d\phi(\gamma(1))(N) + \sum_{\alpha=1}^{m}\frac{df_\alpha(\gamma(1))}{-f_\alpha(\gamma(1))}(N),$$

and by the proof of Theorem 2.2 we estimate

$$| d\phi(\gamma(1))(N) | \leq \frac{k\| N \|^2_H}{1-k\| N \|_H}.$$

$$\left|\sum_{\alpha=1}^{m}\frac{df_\alpha(\gamma(1))}{-f(\gamma(1))}(N)\right| \leq \sqrt{m}\| N \|_{H(\gamma(1))} \leq \sqrt{m}\frac{k\| N \|_H}{1-k\| N \|_H}.$$

Denoting $\rho_k = \frac{1+k\|N\|_H^2}{1-k\|N\|_H}$, it follows

$$|f_0(\gamma(1)) - f_0(x)| \le \mu\sqrt{m}\|N\|_H\rho_k(x).$$

Using the sequence $\{x_i\}$ generated by Newton method, with x as starting point, we get

$$|f_0(\gamma(1)) - f_0(x)| = \left|\sum_{i=1}^{\infty}(f_0(x_{i+1}) - f_0(x_i))\right| \le \sum_{i=1}^{\infty}|f_0(x_{i+1}) - f(x_i)| \le$$

$$\le \sum_{i=1}^{\infty}\|N(x_i,\mu)\|_{H(x_i,\mu)}\rho_k(x_i)\mu\sqrt{m} \le \mu\sqrt{m}\rho_k(x)\sum_{i=1}^{\infty}\|N(x_i,\mu)\|_{H(x_i,\mu)} \le$$

$$\le \mu\sqrt{m}\rho_k(x)\sum_{i=1}^{\infty}\left(\frac{9}{4}k\right)^{2^i-1}\|N\|_H^{2^i} \le \mu\sqrt{m}\rho_k(x)\frac{\|N\|_H}{1-\frac{9}{4}k\|N\|_H}.$$

3. Complexity analysis of the logarithmic barrier algorithm. We look for upper bounds for the total number of outer and inner iterations for the logarithmic barrier algorithm (LBA) applied to the program (P). According the book [1], the analysis refers to long - , medium - , and short - step variants with $\tau = \frac{1}{3k}$.

3.1. Theorem. *Suppose $x_*, z_* = f_0(x_*)$ is the solution of the program (P). After at most*

$$\frac{1}{\theta}\ln\frac{4m\mu_0}{\epsilon}$$

outer iterations, the LBA ends up with an approximate solution x, of the program (P), satisfying $z_* - f_0(x) \le \epsilon$.

Proof. The criteria of stopping is

$$\mu_k = (1-\theta)^k\cdot\mu_0 \le \frac{\epsilon}{4m}, \quad \theta\in(0,1),$$

or

$$-k\ln(1-\theta) \ge \ln\frac{4m\mu_0}{\epsilon}.$$

Since $\sum_{i=1}^{n}\frac{\theta^i}{k} < -\ln(1-\theta)$, we obtain

$$k \ge \frac{1}{\sum_{i=1}^{n}\frac{\theta^i}{i}}\ln\frac{4m\mu_0}{\epsilon}.$$

For simplicity we impose $n = 1$. Also we have

$$z_* - f_0(x(\mu_k)) \le m\mu_k \quad \text{(see [7])}$$

$$f_0(x(\mu_k)) - f_0(x) \le \frac{\frac{1}{3k}}{1-\frac{9}{4}k\cdot\frac{1}{3k}}\cdot\frac{1+k\left(\frac{1}{3k}\right)^2}{1-k\left(\frac{1}{3k}\right)}\mu\sqrt{m},$$

(see Th.2.4) which imply

$$z_* - f_0(x) = z_A - f_0(x(\mu_k)) + f_0(x(\mu_k)) - f_0(x) \le$$

$$\le \mu_k(m+3\sqrt{m}) \le \frac{\epsilon}{4m}(m+3\sqrt{m}) \le \epsilon.$$

Let us prove that Newton method assures a sufficient decrease of the value of logarithmic barrier function.

3.2. Lemma. *If $\gamma(t) = \exp_x(tN), t\in[0,1]$ and $\bar{t} = \frac{1}{1+k\|N\|_H}$, then*

$$\phi(x,\mu) - \phi(\gamma(\bar{t},\mu)) \ge \frac{1}{k^2}(k\|N\|_H - \ln(1+k\|N\|_H)).$$

Proof. Denoting

$$\psi(t) = \phi(x,\mu) - \phi(\gamma(t),\mu),$$

$$\varphi(t) = d\phi(\gamma(t))(N) - (1-t)df(x)(N),$$

it follows

$$\frac{d}{dt}\varphi = -d\varphi(\gamma(t))(N) = -\varphi(t) - (1-t)d\phi(x)(N) = -\varphi(t) + (1-t)\|N\|_H^2 \ge$$

$$\ge (1-t)\|N\|_H^2 - k\|N\|_H^3\frac{t^2}{1-tk\|N\|_H} = v(t).$$

Consequently

$$\psi(t) = \psi(t) - \psi(0) = \int_0^t\frac{d}{ds}\varphi(s)ds \ge \int_0^t v(s)ds =$$

$$= \frac{1}{k^2}(k^2t\|N\|_H^2 + kt\|N\|_H + \ln(1-tk\|N\|_H)) = u(t).$$

The maximum of the function $u : [0, 1] \to R$ is $u(\bar{t})$, where $\bar{t} = \frac{1}{1+k\|N\|_H}$. Substituting $\bar{t}$ into the preceding inequality, we obtain the result of the theorem.

3.3. Theorem. *Each outer iteration requires at most*

$$\frac{22\theta}{(1-\theta)^2}\left(\theta k^2 m + \frac{5}{2}k\sqrt{m}\right) + \frac{22}{3}$$

inner iterations.

Proof. Notations: $\tilde{\mu}$ = the barrier parameter value in an arbitrary outer iteration; $\bar{\mu}$ = the barrier parameter value in the previous outer iteration, x = the iterate at the beginning of the outer iteration. Hence x is centered with respect to $x(\bar{\mu})$ and $\tilde{\mu} = (1-\theta)\bar{\mu}$.

By Lemma 3.2, at each inner iteration, the decrease of the barrier function is at least

$$\frac{1}{k^2}(k\| N \|_H - \ln(1 + k\| N \|_H)).$$

Since this function is increasing in $\| N \|_H$ and $\| N \|_H \geq \frac{1}{3k}$, during each iteration, it follows that the decrease is at least

$$\frac{1}{k^2}\left(\frac{1}{3} - \ln\left(1 + \frac{1}{3}\right)\right) > \frac{1}{22k^2}.$$

If n is the number of inner iterations during one outer iteration, then

$$\frac{n}{22k^2} \leq \varphi(x, \tilde{\mu}) - \varphi(x(\tilde{\mu}), \tilde{\mu}) = \psi(x, \tilde{\mu}).$$

On the other hand, the Meanvalue theorem

$$\psi(x, \tilde{\mu}) = \psi(x, \bar{\mu}) + \frac{d\psi}{d\mu}(x, \mu)\,|_{\mu=\hat{\mu}}\,(\tilde{\mu} - \bar{\mu}), \hat{\mu} \in (\tilde{\mu}, \bar{\mu})$$

and

$$-\frac{d\psi}{d\mu}(x, \mu)\,|_{\mu=\hat{\mu}} = \frac{f_0(x(\mu)) - f_0(x)}{\mu^2}\,|_{\mu=\hat{\mu}} \leq \frac{| f_0(x(\tilde{\mu})) - f_0(x) |}{\tilde{\mu}^2}$$

(see the monotonicity of the objective function along the central path) give

$$\psi(x, \tilde{\mu}) \leq \psi(x, \bar{\mu}) + \frac{| f_0(x(\tilde{\mu})) - f_0(x) |}{\tilde{\mu}^2}(\bar{\mu} - \tilde{\mu}) \leq$$

$$\leq \psi(x, \bar{\mu}) + \left(\frac{| f_0(x(\bar{\mu})) - f_0(x) |}{\tilde{\mu}} + \frac{| f_0(x(\tilde{\mu})) - f_0(x(\bar{\mu})) |}{\tilde{\mu}}\right)\frac{\bar{\mu} - \tilde{\mu}}{\tilde{\mu}}.$$

Since $\| N(x, \bar{\mu}) \|_{H(x,\bar{\mu})} \leq \frac{1}{3k}$, we have $\psi(x, \bar{\mu}) \leq \frac{1}{3k^2}$ and consequently

$$| f_0(x(\bar{\mu})) - f_0(x) | \leq \frac{5}{2}\cdot\frac{\bar{\mu}\sqrt{m}}{k}.$$

Adding, we find

$$f_0(x(\tilde{\mu})) - f_0(x(\bar{\mu})) \leq \sum_{\alpha=1}^{m} y_\alpha(\tilde{\mu}) f_\alpha(x(\tilde{\mu})) -$$

$$-\sum_{\alpha=1}^{m} y_\alpha(\bar{\mu}) f_\alpha(x(\bar{\mu})) = m(\bar{\mu} - \tilde{\mu}) = \theta m \bar{\mu}.$$

We obtain

$$\psi(x, \tilde{\mu}) \leq \frac{1}{3k^2} + \left(\frac{5\sqrt{m}}{2k(1-\theta)} + \frac{\theta m}{1-\theta}\right)\frac{\theta}{1-\theta} = \frac{1}{3k^2} + \frac{\theta}{(1-\theta)^2}\left(\frac{5\sqrt{m}}{2k} + \theta m\right)$$

and hence

$$\frac{n}{22k^2} \leq \frac{1}{3k^2} + \frac{\theta}{(1-\theta)^2}\left(\frac{5\sqrt{m}}{2k} + \theta m\right).$$

Combining the result in the theorems 3.1 and 3.3 we obtain the total number of iterations. More precisely, an upper bound for the total number of Newton iterations is given by

$$\left(\frac{22}{(1-\theta)^2}\left(\theta k^2 m + \frac{5}{2}k\sqrt{m}\right) + \frac{22}{3\theta}\right)\ln\frac{4m\mu_0}{\epsilon}.$$

Therefore, in order to produce an ϵ-optional solution, the algorithm needs:

$O(k^2 m \ln\frac{m\mu_0}{\epsilon})$ Newton iterations for long-step variant $(0 < \theta < 1)$;

$O(k^2\sqrt{m}\ln\frac{m\mu_0}{\epsilon})$ Newton iterations for medium-step variant $\theta = \frac{\nu}{\sqrt{m}} > 0$;

$O(k\sqrt{m}\ln\frac{m\mu_0}{\epsilon})$ Newton iterations for $\theta = \frac{\nu}{k\sqrt{m}} > 0$ (special case of a medium-step variant).

If we want to analyze short-step variants, we can use the next Lemma which show the norm changes when the barrier parameter μ is reduced.

3.4. Lemma. *If* $\mu_+ = (1-\theta)\mu$, *then*

$$\| N(x,\mu_+) \|_{H(x,\mu_+)} \leq \frac{1}{1-\theta}(\| N \|_H + \theta\sqrt{m}).$$

Proof. Obviously $H_+ = H(x,\mu_+) = H - \frac{\theta}{\mu_+}$ Hess $f_0(x)$. By the convexity of $-f_0$, we have $(H_+)^{-1} \leq H^{-1}$. Consequently

$$\| N_+ \|_{H_+} \leq \| N_+ \|_H = \left\| -\frac{df_0(x)}{\mu_+} + \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H =$$

$$= \left\| \frac{1}{1-\theta}\left(-\frac{df_0(x)}{\mu} + \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)}\right) - \frac{\theta}{1-\theta}\sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H \leq$$

$$\leq \frac{1}{1-\theta}\left(\| N \|_H + \theta\left\| \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H\right) \leq \frac{1}{1-\theta}(\| N \|_H + \theta\sqrt{m}),$$

since

$$\left\| \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H \leq \sqrt{m}.$$

Indeed

$$\left\| \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H^2 = H^{-1}\left(\sum_{\beta=1}^{m} \frac{df_\beta(x)}{-f_\beta(x)}, \sum_{\gamma=1}^{m} \frac{df_\gamma(x)}{-f_\gamma(x)}\right)$$

or in the matrix language

$$e^T J^T \left(-\frac{\text{Hess} f_0(x)}{\mu} + \sum_{\alpha=1}^{m} \frac{\text{Hess} f_\alpha(x)}{-f_\alpha(x)} + JJ^T\right)^{-1} Je,$$

where J is the matrix of type $n \times m$ whose columns are $\frac{df_\alpha(x)}{-f_\alpha(x)}$ and $e^T = (1,1,\dots,1)$.

Let A be the positive semidefinite matrix attached to

$$-\frac{\text{Hess} f_0(x)}{\mu} + \sum_{\alpha=1}^{m} \frac{\text{Hess} f_\alpha(x)}{-f_\alpha(x)}.$$

The proper values of the matrix $P = J^T(A+JJ^T)^{-1}J$ are at most 1, since

$$P^2 = J^T(A+JJ^T)^{-1}JJ^T(A+JJ^T)^{-1}J \leq$$

$$\leq J^T(A+JJ^T)^{-1}(A+JJ^T)(A+JJ^T)^{-1}J \leq J^T(A+JJ^T)^{-1}J = \text{P}.$$

It follows

$$\left\| \sum_{\alpha=1}^{m} \frac{df_\alpha(x)}{-f_\alpha(x)} \right\|_H^2 \leq m.$$

3.5. Theorem. *Let* $\gamma(t) = \exp_x(tN)$, $t \in [0,1]$, $x_+ = \gamma(1)$, $\mu_+ = (1-\theta)\mu$, $\theta = \frac{1}{30k\sqrt{m}}$. *If* $\| N \|_H \leq \frac{1}{3k}$, *then*

$$\| N(x_+,\mu_+) \|_{H(x_+,\mu_+)} \leq \frac{1}{3k}.$$

Proof. The Lemmas 3.2, 3.4 imply

$$\| N(x_+,\mu_+) \|_{H(x_+,\mu_+)} \leq \frac{4}{9}k\| N(x_+,\mu_+) \|^2_{H(x_+,\mu_+)} \leq$$

$$\leq \frac{9}{4}k\left(\frac{1}{1-\frac{1}{30k\sqrt{m}}}\frac{1}{3k} + \frac{1}{30k}\right)^2 < \frac{1}{3k}.$$

This last theorem says that for enough small θ, one unit Newton step is sufficient to reach the vicinity of $x(\mu_+)$. The short-step algorithm requires $\mathcal{O}(k\sqrt{m}\ln\frac{m\mu_0}{\epsilon})$ Newton iterations.

REFERENCES

1. H e r t o g, D. den — *Interior point approach to linear, quadratic and convex programming*. Mathematics and Its Applications, 277, Kluwer Academic Publishers, 1994.

2. N e s t e r o v, Y. and N e m i r o v s k y, A. — *Interior-point polynomial algorithms in convex programming*, Siam, 1994.

3. U d r i ş t e, C. — *Kuhn-Tucker theorem on Riemannian manifolds*, Kolloquia Math.Soc. Janos Bolyai, Topics in Diff. Geometry, Debrecen, Hungary (1984), 1247-1259.

4. U d r i ş t e, C. — *Convergence of minimization methods on Riemannian manifolds*, International Workshop on Differential Geometry and Its Applications, Bucharest, July 25-30, Scientific Bulletin, Polytehnica University of Bucharest, Applied Mathematics and Physics, 55, 3-4 (1993), 247-254.
5. U d r i ş t e, C. — *Minimization algorithms on Riemannian manifolds*, Proceedings of the $23-rd$ Conference on Geometry and Topology, Cluj- Napoca, Romania, 1993, 185-193.
6. U d r i ş t e, C. — *Convex functions and optimization methods on Riemannian manifolds*, Mathematics and Its Applications, 297, Kluwer Academic Publishers, 1994.
7. U d r i ş t e, C. — *Optimization methods on Riemannian manifolds*, Proceedings of IRB International Workshop, Monteroduni, Italy, August 8-12, 1995 (in press).
8. U d r i ş t e, C. — *Riemannian convexity in programming (II)*. The Second International Workshop on Differential Geometry and Its Applications, Ovidius University of Constantza, Romania, September 25-28, 1995.
9. U d r i ş t e, C. — *Riemannian convexity*, Proceedings of Second International Workshop on Differential Geometry and Its Applications, Ovidius University of Constantza, Romania, September 25-28, 1995 (in press).

Received : 26.II.1996

University POLITEHNICA of Bucharest
Department of Mathematics I
Splaiul Independentei 313
77206 Bucharest, ROMANIA
Fax: (401) 312.53.65
e-mail:udriste@mathem.pub.ro

ANALELE ŞTIINŢIFICE ALE UNIVERSITĂŢII "AL. I. CUZA" IAŞI
Tomul XLII, Supliment, s.I.a, Matematică, 1996

ENERGY IN RIEMANNIAN GEOMETRY

BY

CHRIS WOOD

Introduction. The use of energy to identify geometric objects with "nice" properties is nothing new; familiar examples include geodesics, harmonic differential forms, harmonic maps, and more recently Yang–Mills connections. It is the purpose of this paper to describe some new ways of bringing energy into Riemannian geometry, within a unified framework which encompasses many different types of geometric structure. Just about everything we have to say takes place in the smooth (C^∞) category, and will be devoid of all but the simplest details, which may be found elsewhere. We begin with a simple example which illustrates the basic idea.

1. The energy of a unit vector field. A vector field X on a manifold M is a section of the tangent bundle TM; that is, a mapping $X: M \to TM$ which is a right inverse for the natural projection $\pi: TM \to M$. If M has a Riemannian metric g, then TM acquires a compatible metric h, called the Sasaki metric (by "compatible" we mean that $\pi: (TM, h) \to (M, g)$ is a Riemannian submersion). It is therefore possible to form the **energy** of X (assuming for convenience that M is compact, and without boundary):

$$E(X) = \frac{1}{2}\int_M |\, dX \,|^2 \, dx$$

where dX is the derivative of the mapping X and dx is the Riemannian volume element of (M, g). We say that X is a **harmonic section of** TM if X is a critical point for E with respect to variations through nearby vector fields, a natural generalization of X being a harmonic map. Now this is a linear variational problem, and may be solved as follows. If

$$dX = d^v X + d^h X$$

Communicated at the 25-th National Conference on Geometry and Topology, September 18-23, 1995, Iaşi, România