TOPONOGOV'S THEOREM FOR METRIC SPACES

by

Conrad Plaut

Max–Planck–Institut
für Mathematik
Gottfried–Claren–Str. 26
5300 Bonn 3
Federal Republic of Germany

MPI/90-47

Toponogov's Theorem for Metric Spaces

by Conrad Plaut

An immediate corollary of the Rauch Comparison Theorem is that if M is a complete Riemannian manifold having sectional curvature $\geq k$, then the Alexandroff distance/angle comparisons are valid in a neighborhood of any point. Toponogov's Theorem states that, in fact, these comparisons are valid *globally*. We show that Toponogov's Theorem is actually true in the most general class of spaces for which the theorem, in its present form, even makes sense: the class of inner metric spaces with (metric) curvature locally bounded below. In the Riemannian special case, the present proof is, to the author's knowlege, the first entirely "metric" proof of Toponogov's Theorem. In particular, it shows for the first time that the powerful improvement from local comparisons to global comparisons is not a differential geometric phenomenon.

Throughout this paper, X will be a metrically complete, locally compact inner metric space. For basic definitions and results, see [R], [P1], or [P2]. We denote by S_p the space of directions at a point p (which in the Riemannian case is simply the unit sphere in T_p), with the angle metric denoted by α. Note that metric completeness and local compactness imply $\{y \in S_p : c(y) \geq \delta\}$ is compact for any positive δ, where c is the cut radius map. We let S_k denote the 2-dimensional, simply connected space form of curvature k. By *monotonicity* we will mean the well-known fact that the angle between two minimal curves of fixed length in S_k is a monotone increasing function of the distance between their non-coincident endpoints. We say that a (geodesic) wedge (γ_1, γ_2) or triangle $(\gamma_1, \gamma_2, \gamma_3)$ is *proper* if γ_1 and γ_3 are minimal and $L(\gamma_2) \le \pi/\sqrt{k}$;

Definition. We say that a triangle $(\gamma_1, \gamma_2, \gamma_3)$ in X is *A1* if there exists a representative triangle $(\bar\gamma_1, \bar\gamma_2, \bar\gamma_3)$ in S_k (i.e., the sides $\bar\gamma_i$ are minimal with $L(\bar\gamma_i) = L(\gamma_i)$), and $\alpha(\bar\gamma_i, \bar\gamma_2) \le \alpha(\gamma_i, \gamma_2)$ for $i = 1, 3$. We say that a wedge $(\gamma_{ab}, \beta_{ac})$ is *A2* if there is a representative wedge $(\bar\gamma_{AB}, \bar\beta_{AC})$ in S_k (i.e., whose sides are minimal with $L(\bar\gamma_{AB}) = L(\gamma_{ab})$, $L(\bar\beta_{AC}) = L(\beta_{ac})$, $\alpha(\bar\gamma_{AB}, \bar\beta_{AC}) = \alpha(\gamma_{ab}, \beta_{ac})$) and $d(B, C) \ge d(b, c)$.

If X has curvature (locally) $\ge$ k, then in a neighborhood of each point every proper triangle is A1 and every proper wedge is A2 (cf. Lemma 1 and [P1]). We prove

Theorem A. *If X is of curvature (locally) $\ge k$, then every proper triangle in X is A1 and every proper wedge in X is A2.*

Corollary B. *If X is of curvature (locally) $\ge k > 0$, then X is compact, with diameter $\le \pi/\sqrt{k}$.*

We say X is *almost Riemannian* ([P2]) of curvature $\ge$ k if X is finite dimensional, geodesically complete, and has curvature (locally) $\geq$ k. Using Theorem A we prove the following precompactness theorem (cf. [GLP]).

Theorem C. *For fixed k, n, and $D > 0$, the class of n-dimensional almost Riemannian spaces of curvature $\geq k$ and diameter $\leq D$ is precompact with the Gromov-Hausdorff metric.*

For the remainder of this paper we assume that X has curvature (locally) $\geq$ k and is at least 2-dimensional (the one dimensional case is trivial). The following lemma formulates a standard technique in proofs of Toponogov's theorem in the Riemannian case (see, e.g., [CE] for an argument).

Lemma 1. *Let $\gamma_{ab} : [0, 1] \to X$ be a geodesic with $L(\gamma_{ab}) \leq \pi/\sqrt{k}$, γ_{ac} be minimal, and $0 = t_0 < t_1 < \ldots < t_i = 1$. Let γ_j denote γ_{ab} restricted to $[t_j, t_{j+1}]$, and suppose α_j is minimal from c to t_j, with $\alpha_0 = \gamma_{ac}$. Then if the triangles $(\alpha_j, \gamma_j, \alpha_{j+1})$ are A1 for $0 \leq j < i$, $(\gamma_{ab}, \gamma_{ac})$ is A2.*

Lemma 2. *For any $\alpha, \beta \in S_p$, $\delta > 0$ and $a_1, a_2 > 0$ such that $a_1 + a_2 = \alpha(\alpha, \beta)$, there exists $\gamma \in S_p$ such that $|\alpha(\alpha, \gamma) - a_1| < \delta$ and $|\alpha(\alpha, \gamma) - a_2| < \delta$.*

Proof. Assume first that $c = \alpha(\alpha, \beta) < \pi$. We need only consider the case $a_1 = a_2 = c/2$. Let $\eta_i : [0, 1] \to X$ be minimal from $\alpha(2^{-i})$ to $\beta(2^{-i})$ and γ_i be minimal from p to $\eta_i(1/2)$; we denote by α_i the restriction of α to $[0, 2^{-i}]$, with similar notation for β. Let $a = \lim_{i\to\infty} \alpha(\alpha, \eta_i)$ and $b = \lim_{i\to\infty} \alpha(\beta, \eta_i)$. By the triangle inequality, $a + b \geq c$. Let $(\bar{\alpha}_i, \bar{\nu}_i, \bar{\gamma}_i)$ and $(\bar{\gamma}_i, \bar{\mu}, \bar{\beta}_i)$ represent $(\alpha_i, \eta_i|_{[0,1/2]}, \gamma_i)$ and $(\gamma_i, \eta_i|_{[1/2,1]}, \beta_i)$, respectively, so that α_i and β_i do not coincide (all of these curves are assumed parameterized on $[0, 1]$). By Lemma 3, [P2], $a = \lim_{i\to\infty} \alpha(\bar{\alpha}_i, \bar{\gamma}_i) = \lim_{i\to\infty} \alpha(\bar{\alpha}_i, \bar{\beta}_i)/2$ and $b = \lim_{i\to\infty} \alpha(\bar{\beta}_i, \bar{\gamma}_i) = \lim_{i\to\infty} \alpha(\bar{\alpha}_i, \bar{\beta}_i)/2$. On the other hand, $d(\bar{\alpha}_i(1), \bar{\beta}_i(1)) \leq d(\alpha(2^{-i}), \eta_i(1/2)) + d(\beta(2^{-i}), \eta_i(1/2))$ and so $\alpha(\alpha, \beta) \leq \lim_{i\to\infty} \alpha(\bar{\alpha}_i, \bar{\beta}_i)$, and the case $c < \pi$ follows.

If $c = \pi$ we can choose a direction distinct from α and β and apply the above special case. □

Remark. One of the few simplifications of the proof of Theorem A in the Riemannian case is that, since a Riemannian manifold has *positive* cut radius and Euclidean tangent space, Lemma 2 is true for $\delta = 0$.

Lemma 3. *If γ_{ab} is minimal then for any $L < L(\gamma_{ab})$ and $\epsilon > 0$ there exists a $\delta > 0$ so that if γ_{ac} is minimal with $L/2 \leq L(\gamma_{ac}) \leq L$ and $d(c, \gamma_{ab}) < \delta$ then $\alpha(\gamma_{ab}, \gamma_{ac}) < \epsilon$.*

Proof. We assume γ_{ab} is unit. Suppose there exist minimal $\gamma_i : [0, 1] \to X$ starting at a such that $d(\gamma_i(1), \gamma_{ab}) \to 0$ and $\alpha(\gamma_i, \gamma_{ab}) \geq \epsilon$. Choosing a subsequence if necessary we can assume $\gamma_i(1) \to \gamma_{ab}(t)$, for some $t \in [L/2, L]$. But then a subsequence of $\{\gamma_i\}$ converges to a minimal curve γ from a to $\gamma(t)$ such that $\alpha(\gamma, \gamma_{ab}) \geq \epsilon > 0$, which contradicts the minimality of γ_{ab}. □

Lemma 4. *Suppose $\alpha, \beta : [o, 1] \to X$ are minimal starting at p with $L(\alpha) \leq \pi/\sqrt{k}$, $L(\beta) < \pi/\sqrt{k}$, and $0 < a = \alpha(\alpha, \beta) < \pi$. Suppose also that $\bar{\alpha}, \bar{\beta} : [0, 1] \to S_k$ are minimal and $(\bar{\alpha}, \bar{\beta})$ represents (α, β). Let $a_1, a_2 > 0$ satisfy $a_1 + a_2 = \alpha(\alpha, \beta)$, $\bar{\gamma}$ be minimal from $\bar{\alpha}(1)$ to $\bar{\beta}(1)$, and t be such that if $\bar{\nu}$ is minimal from $\bar{\alpha}(0)$ to $\bar{\gamma}(t)$ then $\alpha(\bar{\nu}, \bar{\alpha}) = a_1$. If for every $\delta > 0$ there is a geodesic μ starting at p with $L(\mu) = L(\bar{\nu})$ so that $|\alpha(\mu, \alpha) - a_1| < \delta$, $|\alpha(\nu, \alpha) - a_2| < \delta$, and both (α, μ) and (β, μ) are A2, then (α, β) is A2.*

Proof. *Let $\varsigma > 0$* be arbitrary. For sufficiently small δ, there is a representative $(\bar{\alpha}, \bar{\mu})$ of (α, μ) such that $d(\bar{\mu}(1)), \bar{\gamma}(t)) \leq \varsigma$. We assume both $\bar{\mu}$ and μ are parameterized on $[0, 1]$; by A2 and the triangle inequality $d(\alpha(1), \mu(1)) \leq d(\bar{\gamma}(0), \bar{\gamma}(t)) + \varsigma$. Since a similar argument applies to $d(\beta(1), \mu(1))$, we have

$$\begin{aligned} d(\alpha(1), \beta(1)) &\leq d(\alpha(1), \mu(1)) + d(\beta(1), \mu(1)) \\ &\leq d(\bar{\alpha}(1), \bar{\gamma}(t)) + d(\bar{\beta}(1), \bar{\gamma}(t)) + 2\varsigma \\ &= d(\bar{\alpha}(1), \bar{\beta}(1)) + 2\varsigma. \end{aligned}$$ □

Lemma 5. *Given k and $0 < D < \pi/\sqrt{k}$, there exists a $\chi > 0$ so that if $\bar{\gamma}_{AB}$, $\bar{\gamma}_{AC}$ are unit minimal in S_k with $0 < \alpha(\bar{\gamma}_{AB}, \bar{\gamma}_{AC}) < \pi$, $L(\bar{\gamma}_{AB}) \leq D$, and $d(B, C) \leq 3\chi$, then for any $0 < t \leq \min \{L(\bar{\gamma}_{AB}), L(\bar{\gamma}_{AC})\}$ and minimal curve $\bar{\alpha}$ from $\bar{\gamma}_{AB}(t)$ to $\bar{\gamma}_{AC}(t)$, $\max \{d(A, \bar{\alpha}(s)\} < t + \chi$.*

Proof. Since metric balls are convex for $k \leq 0$, we need only consider $k > 0$; by scaling the metric we reduce to $k = 1$, and clearly now we can assume $t > \pi/2$. Let $\chi > 0$ be small enough that $\cos D - (\cos (1.5\chi))(\cos (D+\chi)) > 0$. We fix curves $\bar{\gamma}_{AB}$, $\bar{\gamma}_{AC}$ as above, assume $\bar{\alpha}$ is parameterized on $[0, 1]$ and let $\tau = d(A, \bar{\alpha}(1/2)) = \max \{d(A, \gamma(s)\}$. Letting $\lambda = L(\bar{\alpha})$ and applying the Cosine Law to $\alpha(\bar{\gamma}_{AB}, \bar{\alpha})$ we obtain

$$\frac{\cos \tau - (\cos t)(\cos \lambda/2)}{\sin \lambda/2} = \frac{\cos t - (\cos t)(\cos \lambda)}{\sin \lambda}$$

which reduces to $\cos \tau = \cos t / \cos \lambda/2$.

Applying the sum formula to $\cos (\tau\text{-}t)$ we see that $\tau\text{-}t$ is maximized when $d(A, B) = d(A, C) = t = D$ and $\lambda = 3\chi$. Thus we only need to prove $\cos^{-1}(\cos D / \cos (1.5\chi)) \leq \cos (D+\chi)$, and this follows from the way χ was chosen. □

For $0 < D < \pi/\sqrt{k}$, fix a closed ball $B = \bar{B}(p, D) \subset X$ and a cover U of $\bar{B}(p, 2D)$ by regions of curvature $\geq k$, and let $\chi(U) < D$ be as in Lemma 5 and also less than one eighth of a Lebesque number of U. Let $\tau(U)$ small enough that if $c\bar{\alpha}$, $\bar{\gamma}$ are unit geodesics in S_k with $\alpha(\bar{\alpha}, \bar{\gamma}) \le \tau(U)$, then for all $0 \le t \le D$, $d(\bar{\alpha}(t), \bar{\gamma}(t)) \le \chi(U)$. If $\alpha, \beta : [0, 1] \to B$ are minimal curves starting at p, we call a triangle (α, γ, β) U-*tapered* if $L(\gamma) \le \chi(U)$. We say (α, γ, β) is U-*thin* if $\alpha(\alpha, \beta) \le \tau(U)$ and γ is minimal. We do not require that γ lie in B in either definition, but $\chi(U) < D$ implies γ lies in B(p, 2D).

Note that if $Y \subset B$ has diameter $< 4\chi(U)$ then there exists a region U of curvature $\ge k$ such that every minimal curve joining points in Y lies in U.

Consider the following statements:

S1(n,m). *If (α, γ, β) is* U-*thin such that $L(\alpha) < n \cdot \chi(U)$ and $L(\beta) < m \cdot \chi(U)$, then (α, γ, β) is A1.*

S2(n,m). *If (α, γ, β) is* U-*thin such that $L(\alpha) < n \cdot \chi(U)$ and $L(\beta) < m \cdot \chi(U)$, then (α, β) is A2.*

S3(n). *If (α, γ, β) is* U-*tapered such that $L(\alpha)$, $L(\beta) < n \cdot \chi(U)$, then (α, γ, β) is A1.*

Note that by monotonicity S1(n,m) and S3(n) state equivalently that (α, γ) and (β, γ) are A2. S1(3,3), S2(3,3), and S3(3) are true by the $\chi(U)$ was chosen. We proceed by induction:

Step 1. *S1(n,n) and S2(n,n) imply S2(n, n+1).*

Proof. Fix a U-thin triangle (α, γ, β) such that $n\cdot\chi(U) \le L(\alpha) < (n+1)\cdot\chi(U)$ and $L(\beta) < n\cdot\chi(U)$. Let q lie on α such that $d(p, q) = L(\beta)$, let $x = \alpha(1)$, $y = \beta(1)$ and η be minimal from y to q. If ν is the segment of α from p to q, we obtain from S2(n,n) that (β, ν) is A2 and from S1(n,n) that (ν, η) is A2. S2(n,n) implies dia $\{x, y, q\} \le 2\chi(U)$; if ζ is the segment of α from q to x we have that both (η, ζ) and (ζ, γ) are A2, and that (α, β) is A2 follows from Lemma 1. □

Step 2. *S1(n,n), S2(n,n+1) and S3(n) imply S1(n,n+1).*

Proof. Let (α, γ, β) be as above. The proof that (α, γ) is A2 is similar to the argument in Step 1.

To show (β, γ) is A2 consider an arbitrary $\delta > 0$. Let a be the point on β such that $d(a, y) = \chi(U)$, $R = d(p, a)$, ω denote the segement of β from p to a and ξ be minimal from a to x. Choose a representative $(\bar{\omega}, \bar{\xi})$ in S_k, denoting the corresponding points with capitals. Let $\bar{\mu}$ be unit minimal from P to X and $\bar{\kappa}$ be minimal from A to $\bar{\mu}(R)$. By Lemma 5, for all s, $d(P, \bar{\kappa}(s)) < R + \chi(U) < n\cdot\chi(U)$. Therefore, if δ was chosen sufficiently small and $\kappa : [0, 1] \to X$ is a geodesic starting at a of length $L = L(\bar{\kappa})$ such that $|\alpha(\kappa, \omega) - \alpha(\bar{\kappa}, \bar{\omega})| < \delta$ and $|\alpha(\kappa, \xi) - \alpha(\bar{\kappa}, \bar{\xi})| < \delta$, S3(n) implies that $d(p, \kappa(s)) < n\cdot\chi(U)$ and (κ, ω) is A2. On the other hand, since dia $\{\kappa(1), a, x\}$ is less than $4\chi(U)$, (κ, ξ)

is A2. Lemma 4 now implies (ω, ξ) is A2 and S2(n,n+1) implies (α, ξ, β) is A1. If λ denotes the segment of β from a to y, (ξ, λ, γ) is also A1, and the proof is complete by Lemma 1. □

Step 3. *S1(n,n+1) and S2(n,n+1) imply S1(n+1,n+1) and S2(n+1,n+1).*

Proof. The first implication is a straightforward application of Lemma 1 and the proof of the second is analogous to that of Step 1. □

Step 4. *S3(n), S1(n+1,n+1) and S2(n+1,n+1) imply: If γ is a geodesic in $B(p, (n+1)\chi(B))$, then for any t and minimal α from p to $q = \gamma(t)$ there exists an $\epsilon > 0$ such that if β is minimal from p to $\gamma(s)$ with $|s - t| < \epsilon$, then $(\alpha, \gamma_s, \beta)$ is A1, where γ_s is γ restricted to the interval between s and t.*

Proof. We can assume α is unit of length $L > 4\chi(U)$. Let x be the point on α such that $d(x, q) = \chi(U)$ and denote by ν the segment of α from p to x. By S3(n) there exists an $a_1 > 0$ such that for any geodesic κ starting at x of length $\leq \chi(U)$ with $\alpha(\kappa, \nu) < 2a_1$, (ν, κ) is A2.

By Lemma 3, for any $s > t$ sufficiently close to t and minimal curve ζ from x to $\gamma(s)$, $\alpha(\zeta, \nu)$ is arbitrarily close to π. Let $(\bar{\nu}, \bar{\zeta})$ represent (ν, ζ) in S_k, with $\bar{\zeta}$ parameterized on [0, 1], and let $\bar{\kappa}$ be the geodesic such that $\alpha(\bar{\nu}, \bar{\kappa}) = a_1$ and $\alpha(\bar{\kappa}, \bar{\zeta}) = a_2 = \alpha(\nu, \kappa) - a_1$. For s close enough to t,

$L(\nu) + L(\zeta) < \pi/\sqrt{k}$, and if $\bar{\omega}$ is the unique minimal curve from P to $\bar{\zeta}(1)$, $\bar{\omega}$ intersects $\bar{\kappa}$ at $\bar{\kappa}(r)$, $r \le \chi(U)$. For any $\delta > 0$ we can choose κ as above of length r such that $|\alpha(\kappa, \nu) - a_1| < \delta$ and $|\alpha(\kappa, \zeta) - a_2| < \delta$; applying Lemma 4 we obtain that ν, ζ is A2. If ω is the segment of α from x to q, we have that (ζ, ω) and (ω, γ_s) are both A2, and from S2(n+1,n+1) and Lemma 1 we obtain that (α, γ_s) is A2. Repeating this argument for values $s < t$ we obtain that there exists some ϵ' such that for all s with $|s - t| < \epsilon'$, (α, γ_s) is A2.

To complete the proof we need only show that for any $s_i \to t$ and Cauchy sequence $\{\mu_i\}$, with μ_i minimal from p to $\gamma(s_i)$, (μ_i, γ_i) is A2 for all sufficiently large i (where γ_i denotes the restriction of γ to the interval between t and s_i). Without loss of generality we can assume $s_i > t$. If $\mu = \lim \mu_i$, then μ is minimal from p to q and for all sufficiently large i, (μ, μ_i) is U-thin. The argument is now finished by S1(n+1,n+1). □

Step 5. *S3(n), S1(n+1,n+1) and S2(n+1,n+1) imply S3(n+1,n+1) (and the induction is complete).*

Proof. Let (α, γ, β) be U-tapered with $L(\alpha)$, $L(\beta) < n\cdot\chi(U)$; we assume γ is parameterized on [0, 1]. For $s > 0$, let γ_s denote $\gamma|_{[0,s]}$, and denote by A1(s) the statement: *for every minimal β_s from p to $\gamma(s)$, $(\alpha, \gamma_s, \beta_s)$ is A1.* Step 4 implies that A1(s) is true for sufficiently small s, and Step 4 and Lemma 1 prove immediately that if A1(T) is true for some T, then A1(s) is true for all $s > T$ sufficiently close to T. Likewise, if $A1(s)$ is true for all $s < T$ then $A1(T)$ is true; it follows that $A1(1)$ holds, which is even more than we needed to prove. □

Step 6. *Every proper wedge in $\bar{B}(p, D)$ such that $\alpha(1) = p$ is A2.*

Proof. Let (α, β) be a proper wedge in $\bar{B}(p, D)$ with $\alpha(1) = p$. (We assume both curves are parameterized on $[0, 1]$.) Subdivide β into minimal curves β_i of length $< \chi(U)$. Choosing minimal curves from p to the endpoints of each β_i we obtain U-tapered triangles; applying Lemma 1 completes the proof. □

The proof of Theorem A is complete by Step 6 in the case $k \leq 0$. For the case $k > 0$, note that a limit argument using $S2(n,n)$ shows that if α and β are minimal of length $\pi/\sqrt{k}$ and have one endpoint in common, then they have the other endpoint in common. This proves Corollary B. Furthermore, if (α, γ) is a proper wedge, $d(\alpha(1), \gamma(s)) = \pi/\sqrt{k}$ at at most a finite number of values s. Taking $p = \alpha(1)$ and D sufficiently close to $\pi/\sqrt{k}$ we can now use an argument similar to the proof of Step 4, and Lemma 1, to complete the proof of Theorem A.

Proof of Theorem C. For any compact metric space Y we denote by $N(\epsilon, r, Y)$ the maximum number of disjoint balls of radius ϵ that can be put in a ball of radius r in Y. Suppose $\dim X = n$. By [GLP], Proposition 5.2, it suffices to prove that

$N(\epsilon, r, X) \leq N(\epsilon, r, S^n_L)$, where for simplicity we use $L = \min\{0, k\}$ instead of k. Let B(x, r) be given, and endow $B(0, r) \subset \overline{T}_x$ with the metric δ defined in the proof of Proposition 19, [P2], so that B(0, r) is isometric to a ball of radius r in S^2_k. Let $B_i = B_i(p_i, \epsilon)$ be a collection of N disjoint balls in B(x, r). Let $d_i = d(x, p_i) \leq r - \epsilon$, and $v_i \in S_x$ be minimal from x to p_i. Then since $\exp_x$ is distance decreasing, $\exp_x(B_\delta(v_i, \epsilon)) \subseteq B_i$, and the balls $B_\delta(v_i, \epsilon)$ are N disjoint ϵ-balls in $B(0, r) = B(z, r) \subset S^n_k$. □

The above argument is easily modified to obtain a "pointed" precompactness theorem without an upper bound on the diameter (cf. [GLP]).

Bibliography

[CE] Cheeger, J., and Ebin, D. G. *Comparison Theorems in Riemannian Geometry*, North Holland, Amsterdam, 1975.

[GLP] Gromov, M., Lafontaine, J., and Pansu, P. *Structure Mètrique pour les Variètès Riemanniennes*, Cedic/Fernant Nathan, Paris, 1981.

[K] Karcher, H. *Riemannian Comparison Constructions*, preprint.

[P1] Plaut, C. A metric characterization of manifolds with boundary, Max-Planck-Institut preprint no. 89-60.

[P2] Plaut, C. Almost Riemannian spaces, Max-Planck-Institut preprint no. 90-20.

[R] Rinow, W. *Die Innere Geometrie der Metrischen Räume*, Springer-Verlag, Berlin, 1961.