Gromov hyperbolic spaces in Isabelle

Sebastien Gouezel

Abstract

A geodesic metric space is Gromov hyperbolic if all its geodesic triangles are thin, i.e., every side is contained in a fixed thickening of the two other sides. While this definition looks innocuous, it has proved extremely important and versatile in modern geometry since its introduction by Gromov. We formalize the basic classical properties of Gromov hyperbolic spaces, notably the Morse lemma asserting that quasigeodesics are close to geodesics, the invariance of hyperbolicity under quasi-isometries, we define and study the Gromov boundary and its associated distance, and prove that a quasi-isometry between Gromov hyperbolic spaces extends to a homeomorphism of the boundaries. We also classify the isometries of hyperbolic spaces into elliptic, parabolic and loxodromic ones, both in terms of translation length and of fixed points at infinity. We also prove a less classical theorem, by Bonk and Schramm, asserting that a Gromov hyperbolic space embeds isometrically in a geodesic Gromov-hyperbolic space. As the original proof uses a transfinite sequence of Cauchy completions, this is an interesting formalization exercise. Along the way, we introduce basic material on isometries, quasi-isometries, geodesic spaces, the Hausdorff distance, the Cauchy completion of a metric space, and the exponential on extended real numbers.

Contents

1 Additions to the library **3**
1.1 Mono intros . . . 3
1.2 More topology . . . 6
1.2.1 Homeomorphisms . . . 8
1.2.2 Proper spaces . . . 10
1.2.3 Miscellaneous topology . . . 11
1.2.4 Measure of balls . . . 12
1.2.5 infdist and closest point projection . . . 13
1.3 Material on ereal and ennreal . . . 14
1.4 Miscellaneous . . . 17
1.4.1 Liminfs and Limsups . . . 17
1.4.2 Bounding the cardinality of a finite set . . . 18
1.5 Manipulating finite ordered sets . . . 19
1.6 Well-orders . . . 21

2 The exponential on extended real numbers. **22**

3 Hausdorff distance **30**
3.1 Preliminaries 30
3.2 Hausdorff distance 30

4 Isometries **32**

5 Geodesic spaces **35**
5.1 Geodesic segments in general metric spaces 35
5.2 Geodesic subsets 42
5.3 Geodesic spaces 43
5.4 Uniquely geodesic spaces 43
5.5 A complete metric space with middles is geodesic. 45

6 Quasi-isometries **47**
6.1 Basic properties of quasi-isometries 47
6.2 Quasi-isometric isomorphisms 49
6.3 Quasi-isometries of Euclidean spaces. 51
6.4 Quasi-geodesics 52

7 The metric completion of a metric space **52**
7.1 Definition of the metric completion 53
7.2 Isometric embedding of a space in its metric completion . . . 54
7.3 The metric completion of a second countable space is second countable 55

8 Gromov hyperbolic spaces **56**
8.1 Definition, basic properties 56
8.2 Typeclass for Gromov hyperbolic spaces 58

9 Metric trees **62**

10 Quasiconvexity **65**

11 The Morse-Gromov Theorem **68**

12 The Bonk Schramm extension **76**
12.1 Unfolded Bonk Schramm extension 77
12.2 The Bonk Schramm extension 82

13 Bonk-Schramm extension of hyperbolic spaces **84**
13.1 The Bonk-Schramm extension preserves hyperbolicity 84
13.2 Applications 84

14 Constructing a distance from a quasi-distance **86**

15 The Gromov completion of a hyperbolic space **88**
15.1 The Gromov boundary as a set 88
15.2 Extending the original distance and the original Gromov product to the completion 91
15.3 Construction of the distance on the Gromov completion . . . 94
15.4 Characterizing convergence in the Gromov boundary 97
15.5 Continuity properties of the extended Gromov product and distance 100
15.6 Topology of the Gromov boundary 102
15.7 The Gromov completion of the real line. 104

16 Extension of quasi-isometries to the boundary **105**

17 Extensions of isometries to the boundary **109**

18 Busemann functions **111**

19 Classification of isometries on a Gromov hyperbolic space **114**
19.1 The translation length 115
19.2 The strength of an isometry at a fixed point at infinity 116
19.3 Elliptic isometries 117
19.4 Parabolic and loxodromic isometries 118
19.4.1 Parabolic isometries 121
19.4.2 Loxodromic isometries 122

1 Additions to the library

theory *Library-Complements*
imports *HOL−Analysis.Analysis HOL−Cardinals.Cardinal-Order-Relation*
begin

1.1 Mono intros

We have a lot of (large) inequalities to prove. It is very convenient to have a set of introduction rules for this purpose (a lot should be added to it, I have put here all the ones I needed).

The typical use case is when one wants to prove some inequality, say $\exp(x * x) \leq y + \exp(1 + z * z + y)$, assuming $y \geq 0$ and $0 \leq x \leq z$. One would write it has

```
have "0 + \exp(0 + x * x + 0) < = y + \exp(1 + z * z + y)"
using ‘y > = 0‘ ‘x < = z‘ by (intro mono_intros)
```

When the left and right hand terms are written in completely analogous ways as above, then the introduction rules (that contain monotonicity of addition, of the exponential, and so on) reduce this to comparison of elementary terms in the formula. This is a very naive strategy, that fails in many situations, but that is very efficient when used correctly.

named-theorems *mono-intros structural introduction rules to prove inequalities*
declare *le-imp-neg-le* [*mono-intros*]
declare *add-left-mono* [*mono-intros*]
declare *add-right-mono* [*mono-intros*]
declare *add-strict-left-mono* [*mono-intros*]
declare *add-strict-right-mono* [*mono-intros*]
declare *add-mono* [*mono-intros*]
declare *add-less-le-mono* [*mono-intros*]
declare *diff-right-mono* [*mono-intros*]
declare *diff-left-mono* [*mono-intros*]
declare *diff-mono* [*mono-intros*]
declare *mult-left-mono* [*mono-intros*]
declare *mult-right-mono* [*mono-intros*]
declare *mult-mono* [*mono-intros*]
declare *max.mono* [*mono-intros*]
declare *min.mono* [*mono-intros*]
declare *power-mono* [*mono-intros*]
declare *ln-ge-zero* [*mono-intros*]
declare *ln-le-minus-one* [*mono-intros*]
declare *ennreal-minus-mono* [*mono-intros*]
declare *ennreal-leI* [*mono-intros*]
declare *e2ennreal-mono* [*mono-intros*]
declare *enn2ereal-nonneg* [*mono-intros*]
declare *zero-le* [*mono-intros*]
declare *top-greatest* [*mono-intros*]
declare *bot-least* [*mono-intros*]
declare *dist-triangle* [*mono-intros*]
declare *dist-triangle2* [*mono-intros*]
declare *dist-triangle3* [*mono-intros*]
declare *exp-ge-add-one-self* [*mono-intros*]
declare *exp-gt-one* [*mono-intros*]
declare *exp-less-mono* [*mono-intros*]
declare *dist-triangle* [*mono-intros*]
declare *abs-triangle-ineq* [*mono-intros*]
declare *abs-triangle-ineq2* [*mono-intros*]
declare *abs-triangle-ineq2-sym* [*mono-intros*]
declare *abs-triangle-ineq3* [*mono-intros*]
declare *abs-triangle-ineq4* [*mono-intros*]
declare *Liminf-le-Limsup* [*mono-intros*]
declare *ereal-liminf-add-mono* [*mono-intros*]
declare *le-of-int-ceiling* [*mono-intros*]
declare *ereal-minus-mono* [*mono-intros*]
declare *infdist-triangle* [*mono-intros*]
declare *divide-right-mono* [*mono-intros*]
declare *self-le-power* [*mono-intros*]

lemma *ln-le-cancelI* [*mono-intros*]:
  assumes $(0::real) < x$ $x \leq y$
  shows $ln\ x \leq ln\ y$
⟨*proof*⟩

lemma *exp-le-cancelI* [*mono-intros*]:
  assumes $x \leq (y::real)$
  shows $exp\ x \leq exp\ y$
⟨*proof*⟩

lemma *mult-ge1-mono* [*mono-intros*]:
  assumes $a \geq (0::'a::linordered\text{-}idom)$ $b \geq 1$
  shows $a \leq a * b$ $a \leq b * a$
⟨*proof*⟩

A few convexity inequalities we will need later on.

lemma *xy-le-uxx-vyy* [*mono-intros*]:
  assumes $u > 0$ $u * v = (1::real)$
  shows $x * y \leq u * x\hat{}2/2 + v * y\hat{}2/2$
⟨*proof*⟩

lemma *xy-le-xx-yy* [*mono-intros*]:
  $x * y \leq x\hat{}2/2 + y\hat{}2/2$ **for** $x\ y::real$
⟨*proof*⟩

lemma *ln-squared-bound* [*mono-intros*]:
  $(ln\ x)\hat{}2 \leq 2 * x - 2$ **if** $x \geq 1$ **for** $x::real$
⟨*proof*⟩

In the next lemma, the assumptions are too strong (negative numbers less than −1 also work well to have a square larger than 1), but in practice one proves inequalities with nonnegative numbers, so this version is really the useful one for `mono_intros`.

lemma *mult-ge1-powers* [*mono-intros*]:
  assumes $a \geq (1::'a::linordered\text{-}idom)$
  shows $1 \leq a * a$ $1 \leq a * a * a$ $1 \leq a * a * a * a$
⟨*proof*⟩

lemmas [*mono-intros*] = *ln-bound*

lemma *mono-cSup*:
  fixes $f :: 'a::conditionally\text{-}complete\text{-}lattice \Rightarrow 'b::conditionally\text{-}complete\text{-}lattice$
  assumes *bdd-above A* $A \neq \{\}$ *mono f*
  shows $Sup\ (f`A) \leq f\ (Sup\ A)$
⟨*proof*⟩

lemma *mono-cSup-bij*:
  fixes $f :: 'a::conditionally\text{-}complete\text{-}linorder \Rightarrow 'b::conditionally\text{-}complete\text{-}linorder$
  assumes *bdd-above A* $A \neq \{\}$ *mono f bij f*

shows $Sup\ (f`A) = f(Sup\ A)$
⟨*proof*⟩

1.2 More topology

In situations of interest to us later on, convergence is well controlled only for sequences living in some dense subset of the space (but the limit can be anywhere). This is enough to establish continuity of the function, if the target space is well enough separated.

The statement we give below is very general, as we do not assume that the function is continuous inside the original set S, it will typically only be continuous at a set T contained in the closure of S. In many applications, T will be the closure of S, but we are also thinking of the case where one constructs an extension of a function inside a space, to its boundary, and the behaviour at the boundary is better than inside the space. The example we have in mind is the extension of a quasi-isometry to the boundary of a Gromov hyperbolic space.

In the following criterion, we assume that if u_n inside S converges to a point at the boundary T, then $f(u_n)$ converges (where f is some function inside). Then, we can extend the function f at the boundary, by picking the limit value of $f(u_n)$ for some sequence converging to u_n. Then the lemma asserts that f is continuous at every point b on the boundary.

The proof is done in two steps:

1. First, if v_n is another inside sequence tending to the same point b on the boundary, then $f(v_n)$ converges to the same value as $f(u_n)$: this is proved by considering the sequence w equal to u at even times and to v at odd times, and saying that $f(w_n)$ converges. Its limit is equal to the limit of $f(u_n)$ and of $f(v_n)$, so they have to coincide.

2. Now, consider a general sequence v (in the space or the boundary) converging to b. We want to show that $f(v_n)$ tends to $f(b)$. If v_n is inside S, we have already done it in the first step. If it is on the boundary, on the other hand, we can approximate it by an inside point w_n for which $f(w_n)$ is very close to $f(v_n)$. Then w_n is an inside sequence converging to b, hence $f(w_n)$ converges to $f(b)$ by the first step, and then $f(v_n)$ also converges to $f(b)$. The precise argument is more conveniently written by contradiction. It requires good separation properties of the target space.

First, we introduce the material to interpolate between two sequences, one at even times and the other one at odd times.

definition *even-odd-interpolate*::$(nat \Rightarrow 'a) \Rightarrow (nat \Rightarrow 'a) \Rightarrow (nat \Rightarrow 'a)$
  where *even-odd-interpolate* $u\ v\ n = (if\ even\ n\ then\ u\ (n\ div\ 2)\ else\ v\ (n\ div\ 2))$

lemma *even-odd-interpolate-compose*:
even-odd-interpolate (*f o u*) (*f o v*) = *f o* (*even-odd-interpolate u v*)
⟨*proof*⟩

lemma *even-odd-interpolate-filterlim*:
filterlim u F sequentially ∧ *filterlim v F sequentially* ⟷ *filterlim* (*even-odd-interpolate u v*) *F sequentially*
⟨*proof*⟩

Then, we prove the continuity criterion for extensions of functions to the boundary T of a set S. The first assumption is that $f(u_n)$ converges when f converges to the boundary, and the second one that the extension of f to the boundary has been defined using the limit along some sequence tending to the point under consideration. The following criterion is the most general one, but this is not the version that is most commonly applied so we use a prime in its name.

lemma *continuous-at-extension-sequentially'*:
fixes *f* :: *'a*::{*first-countable-topology*, *t2-space*} ⇒ *'b*::*t3-space*
assumes *b* ∈ *T*
⋀*u b*. (∀ *n*. *u n* ∈ *S*) ⟹ *b* ∈ *T* ⟹ *u* ⟶ *b* ⟹ *convergent* (λ*n*. *f* (*u n*))
⋀*b*. *b* ∈ *T* ⟹ ∃ *u*. (∀ *n*. *u n* ∈ *S*) ∧ *u* ⟶ *b* ∧ ((λ*n*. *f* (*u n*)) ⟶ *f b*)
shows *continuous* (*at b within* (*S* ∪ *T*)) *f*
⟨*proof*⟩

We can specialize the previous statement to the common case where one already knows the sequential continuity of f along sequences in S converging to a point in T. This will be the case in most –but not all– applications. This is a straightforward application of the above criterion.

proposition *continuous-at-extension-sequentially*:
fixes *f* :: *'a*::{*first-countable-topology*, *t2-space*} ⇒ *'b*::*t3-space*
assumes *a* ∈ *T*
T ⊆ *closure S*
⋀*u b*. (∀ *n*. *u n* ∈ *S*) ⟹ *b* ∈ *T* ⟹ *u* ⟶ *b* ⟹ (λ*n*. *f* (*u n*)) ⟶ *f b*
shows *continuous* (*at a within* (*S* ∪ *T*)) *f*
⟨*proof*⟩

We also give global versions. We can only express the continuity on T, so this is slightly weaker than the previous statements since we are not saying anything on inside sequences tending to T – but in cases where T contains S these statements contain all the information.

lemma *continuous-on-extension-sequentially'*:
fixes *f* :: *'a*::{*first-countable-topology*, *t2-space*} ⇒ *'b*::*t3-space*
assumes ⋀*u b*. (∀ *n*. *u n* ∈ *S*) ⟹ *b* ∈ *T* ⟹ *u* ⟶ *b* ⟹ *convergent* (λ*n*. *f* (*u n*))

$\bigwedge b.\ b \in T \Longrightarrow \exists\, u.\ (\forall\, n.\ u\ n \in S) \wedge u \longrightarrow b \wedge ((\lambda n.\ f\ (u\ n)) \longrightarrow f\ b)$
shows *continuous-on T f*
⟨*proof*⟩

lemma *continuous-on-extension-sequentially*:
fixes f :: $'a$::{*first-countable-topology*, *t2-space*} ⇒ $'b$::*t3-space*
assumes $T \subseteq$ *closure* S
$\bigwedge u\ b.\ (\forall\, n.\ u\ n \in S) \Longrightarrow b \in T \Longrightarrow u \longrightarrow b \Longrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ b$
shows *continuous-on T f*
⟨*proof*⟩

1.2.1 Homeomorphisms

A variant around the notion of homeomorphism, which is only expressed in terms of the function and not of its inverse.

definition *homeomorphism-on*::$'a$ *set* ⇒ ($'a$::*topological-space* ⇒ $'b$::*topological-space*) ⇒ *bool*
where *homeomorphism-on S f* = (∃ *g*. *homeomorphism S* (*f*‘*S*) *f g*)

lemma *homeomorphism-on-continuous*:
assumes *homeomorphism-on S f*
shows *continuous-on S f*
⟨*proof*⟩

lemma *homeomorphism-on-bij*:
assumes *homeomorphism-on S f*
shows *bij-betw f S* (*f*‘*S*)
⟨*proof*⟩

lemma *homeomorphism-on-homeomorphic*:
assumes *homeomorphism-on S f*
shows *S homeomorphic* (*f*‘*S*)
⟨*proof*⟩

lemma *homeomorphism-on-compact*:
fixes *f*::$'a$::*topological-space* ⇒ $'b$::*t2-space*
assumes *continuous-on S f*
compact S
inj-on f S
shows *homeomorphism-on S f*
⟨*proof*⟩

lemma *homeomorphism-on-subset*:
assumes *homeomorphism-on S f*
$T \subseteq S$
shows *homeomorphism-on T f*
⟨*proof*⟩

lemma *homeomorphism-on-empty* [*simp*]:
homeomorphism-on {} *f*
⟨*proof*⟩

lemma *homeomorphism-on-cong*:
assumes *homeomorphism-on X f*
$X' = X \bigwedge x.\ x \in X \Longrightarrow f'\ x = f\ x$
shows *homeomorphism-on* $X'\ f'$
⟨*proof*⟩

lemma *homeomorphism-on-inverse*:
fixes *f*::′*a*::*topological-space* ⇒ ′*b*::*topological-space*
assumes *homeomorphism-on X f*
shows *homeomorphism-on* (*f*‘*X*) (*inv-into X f*)
⟨*proof*⟩

Characterization of homeomorphisms in terms of sequences: a map is a homeomorphism if and only if it respects convergent sequences.

lemma *homeomorphism-on-compose*:
assumes *homeomorphism-on S f*
$x \in S$
eventually $(\lambda n.\ u\ n \in S)\ F$
shows $(u \longrightarrow x)\ F \longleftrightarrow ((\lambda n.\ f\ (u\ n)) \longrightarrow f\ x)\ F$
⟨*proof*⟩

lemma *homeomorphism-on-sequentially*:
fixes *f*::′*a*::{*first-countable-topology*, *t2-space*} ⇒ ′*b*::{*first-countable-topology*, *t2-space*}
assumes $\bigwedge x\ u.\ x \in S \Longrightarrow (\forall n.\ u\ n \in S) \Longrightarrow u \longrightarrow x \longleftrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ x$
shows *homeomorphism-on S f*
⟨*proof*⟩

lemma *homeomorphism-on-UNIV-sequentially*:
fixes *f*::′*a*::{*first-countable-topology*, *t2-space*} ⇒ ′*b*::{*first-countable-topology*, *t2-space*}
assumes $\bigwedge x\ u.\ u \longrightarrow x \longleftrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ x$
shows *homeomorphism-on UNIV f*
⟨*proof*⟩

Now, we give similar characterizations in terms of sequences living in a dense subset. As in the sequential continuity criteria above, we first give a very general criterion, where the map does not have to be continuous on the approximating set S, only on the limit set T, without any a priori identification of the limit. Then, we specialize this statement to a less general but often more usable version.

lemma *homeomorphism-on-extension-sequentially-precise*:
fixes *f*::′*a*::{*first-countable-topology*, *t3-space*} ⇒ ′*b*::{*first-countable-topology*, *t3-space*}
assumes $\bigwedge u\ b.\ (\forall n.\ u\ n \in S) \Longrightarrow b \in T \Longrightarrow u \longrightarrow b \Longrightarrow$ *convergent* $(\lambda n.\ f$

$(u\ n))$
$\bigwedge u\ c.\ (\forall n.\ u\ n \in S) \Longrightarrow c \in f`T \Longrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow c \Longrightarrow convergent$
u
$\bigwedge b.\ b \in T \Longrightarrow \exists u.\ (\forall n.\ u\ n \in S) \wedge u \longrightarrow b \wedge ((\lambda n.\ f\ (u\ n)) \longrightarrow$
$f\ b)$
$\bigwedge n.\ u\ n \in S \cup T\ l \in T$
shows $u \longrightarrow l \longleftrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ l$
⟨*proof*⟩

lemma *homeomorphism-on-extension-sequentially'*:
fixes f::$'a$::{*first-countable-topology*, *t3-space*} ⇒ $'b$::{*first-countable-topology*, *t3-space*}
assumes $\bigwedge u\ b.\ (\forall n.\ u\ n \in S) \Longrightarrow b \in T \Longrightarrow u \longrightarrow b \Longrightarrow convergent\ (\lambda n.\ f$
$(u\ n))$
$\bigwedge u\ c.\ (\forall n.\ u\ n \in S) \Longrightarrow c \in f`T \Longrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow c \Longrightarrow convergent$
u
$\bigwedge b.\ b \in T \Longrightarrow \exists u.\ (\forall n.\ u\ n \in S) \wedge u \longrightarrow b \wedge ((\lambda n.\ f\ (u\ n)) \longrightarrow$
$f\ b)$
shows *homeomorphism-on T f*
⟨*proof*⟩

proposition *homeomorphism-on-extension-sequentially*:
fixes f::$'a$::{*first-countable-topology*, *t3-space*} ⇒ $'b$::{*first-countable-topology*, *t3-space*}
assumes $\bigwedge u\ b.\ (\forall n.\ u\ n \in S) \Longrightarrow u \longrightarrow b \longleftrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ b$
$T \subseteq closure\ S$
shows *homeomorphism-on T f*
⟨*proof*⟩

lemma *homeomorphism-on-UNIV-extension-sequentially*:
fixes f::$'a$::{*first-countable-topology*, *t3-space*} ⇒ $'b$::{*first-countable-topology*, *t3-space*}
assumes $\bigwedge u\ b.\ (\forall n.\ u\ n \in S) \Longrightarrow u \longrightarrow b \longleftrightarrow (\lambda n.\ f\ (u\ n)) \longrightarrow f\ b$
$closure\ S = UNIV$
shows *homeomorphism-on UNIV f*
⟨*proof*⟩

1.2.2 Proper spaces

Proper spaces, i.e., spaces in which every closed ball is compact – or, equivalently, any closed bounded set is compact.

definition *proper*::($'a$::*metric-space*) *set* ⇒ *bool*
where *proper S* ≡ ($\forall$ *x r*. *compact* (*cball x r* ∩ *S*))

lemma *properI*:
assumes $\bigwedge x\ r.\ compact\ (cball\ x\ r \cap S)$
shows *proper S*
⟨*proof*⟩

lemma *proper-compact-cball*:
assumes *proper* (*UNIV*::$'a$::*metric-space set*)
shows *compact* (*cball* (x::$'a$) *r*)

⟨*proof*⟩

lemma *proper-compact-bounded-closed*:
assumes *proper* (*UNIV*::*'a*::*metric-space set*) *closed* (*S*::*'a set*) *bounded S*
shows *compact S*
⟨*proof*⟩

lemma *proper-real* [*simp*]:
proper (*UNIV*::*real set*)
⟨*proof*⟩

lemma *complete-of-proper*:
assumes *proper S*
shows *complete S*
⟨*proof*⟩

lemma *proper-of-compact*:
assumes *compact S*
shows *proper S*
⟨*proof*⟩

lemma *proper-Un*:
assumes *proper A proper B*
shows *proper* ($A \cup B$)
⟨*proof*⟩

1.2.3 Miscellaneous topology

When manipulating the triangle inequality, it is very frequent to deal with 4 points (and automation has trouble doing it automatically). Even sometimes with 5 points...

lemma *dist-triangle4* [*mono-intros*]:
dist x t ≤ *dist x y* + *dist y z* + *dist z t*
⟨*proof*⟩

lemma *dist-triangle5* [*mono-intros*]:
dist x u ≤ *dist x y* + *dist y z* + *dist z t* + *dist t u*
⟨*proof*⟩

A thickening of a compact set is closed.

lemma *compact-has-closed-thickening*:
assumes *compact C*
continuous-on C f
shows *closed* (⋃*x*∈*C*. *cball x* (*f x*))
⟨*proof*⟩

congruence rule for continuity. The assumption that $fy = gy$ is necessary since `at x` is the pointed neighborhood at x.

lemma *continuous-within-cong*:
assumes *continuous* (*at y within S*) *f*
eventually (λx. *f x* = *g x*) (*at y within S*)
f y = *g y*
shows *continuous* (*at y within S*) *g*
⟨*proof*⟩

A function which tends to infinity at infinity, on a proper set, realizes its infimum

lemma *continuous-attains-inf-proper*:
fixes *f* :: *'a*::*metric-space* ⇒ *'b*::*linorder-topology*
assumes *proper s a* ∈ *s*
continuous-on s f
⋀*z*. *z* ∈ *s* − *cball a r* ⟹ *f z* ≥ *f a*
shows ∃ *x*∈*s*. ∀ *y*∈*s*. *f x* ≤ *f y*
⟨*proof*⟩

1.2.4 Measure of balls

The image of a ball by an affine map is still a ball, with explicit center and radius. (Now unused)

lemma *affine-image-ball* [*simp*]:
(λy. $R *_R y + x$) ‘ *cball 0 1* = *cball* (*x*::(*'a*::*real-normed-vector*)) |*R*|
⟨*proof*⟩

From the rescaling properties of Lebesgue measure in a euclidean space, it follows that the measure of any ball can be expressed in terms of the measure of the unit ball.

lemma *lebesgue-measure-ball*:
assumes *R* ≥ *0*
shows *measure lborel* (*cball* (*x*::(*'a*::*euclidean-space*)) *R*) = *R*^(*DIM*(*'a*)) ∗ *measure lborel* (*cball* (*0*::*'a*) *1*)
emeasure lborel (*cball* (*x*::(*'a*::*euclidean-space*)) *R*) = *R*^(*DIM*(*'a*)) ∗ *emeasure lborel* (*cball* (*0*::*'a*) *1*)
⟨*proof*⟩

We show that the unit ball has positive measure – this is obvious, but useful. We could show it by arguing that it contains a box, whose measure can be computed, but instead we say that if the measure vanished then the measure of any ball would also vanish, contradicting the fact that the space has infinite measure. This avoids all computations.

lemma *lebesgue-measure-ball-pos*:
emeasure lborel (*cball* (*0*::*'a*::*euclidean-space*) *1*) > *0*
measure lborel (*cball* (*0*::*'a*::*euclidean-space*) *1*) > *0*
⟨*proof*⟩

1.2.5 infdist and closest point projection

The distance to a union of two sets is the minimum of the distance to the two sets.

lemma *infdist-union-min* [*mono-intros*]:
  assumes $A \neq \{\}$ $B \neq \{\}$
  shows $infdist\ x\ (A \cup B) = min\ (infdist\ x\ A)\ (infdist\ x\ B)$
⟨*proof*⟩

The distance to a set is non-increasing with the set.

lemma *infdist-mono* [*mono-intros*]:
  assumes $A \subseteq B$ $A \neq \{\}$
  shows $infdist\ x\ B \leq infdist\ x\ A$
  ⟨*proof*⟩

If a set is proper, then the infimum of the distances to this set is attained.

lemma *infdist-proper-attained*:
  assumes $proper\ C$ $C \neq \{\}$
  shows $\exists\, c{\in}C.\ infdist\ x\ C = dist\ x\ c$
⟨*proof*⟩

lemma *infdist-almost-attained*:
  assumes $infdist\ x\ X < a$ $X \neq \{\}$
  shows $\exists\, y{\in}X.\ dist\ x\ y < a$
⟨*proof*⟩

lemma *infdist-triangle-abs* [*mono-intros*]:
  $|infdist\ x\ A - infdist\ y\ A| \leq dist\ x\ y$
⟨*proof*⟩

The next lemma is missing in the library, contrary to its cousin `continuous_infdist`.

The infimum of the distance to a singleton set is simply the distance to the unique member of the set.

The closest point projection of x on A. It is not unique, so we choose one point realizing the minimal distance. And if there is no such point, then we use x, to make some statements true without any assumption.

definition *proj-set*::$'a$::*metric-space* $\Rightarrow$ $'a\ set \Rightarrow\ 'a\ set$
  where $proj\text{-}set\ x\ A = \{y \in A.\ dist\ x\ y = infdist\ x\ A\}$

definition *distproj*::$'a$::*metric-space* $\Rightarrow$ $'a\ set \Rightarrow\ 'a$
  where $distproj\ x\ A = (if\ proj\text{-}set\ x\ A \neq \{\}\ then\ SOME\ y.\ y \in proj\text{-}set\ x\ A\ else\ x)$

lemma *proj-setD*:
  assumes $y \in proj\text{-}set\ x\ A$
  shows $y \in A$ $dist\ x\ y = infdist\ x\ A$

⟨proof⟩

lemma *proj-setI*:
  assumes $y \in A$ *dist* x $y \leq$ *infdist* x A
  shows $y \in$ *proj-set* x A
⟨proof⟩

lemma *proj-setI'*:
  assumes $y \in A$ $\bigwedge z.\ z \in A \Longrightarrow$ *dist* x $y \leq$ *dist* x z
  shows $y \in$ *proj-set* x A
⟨proof⟩

lemma *distproj-in-proj-set*:
  assumes *proj-set* x $A \neq \{\}$
  shows *distproj* x $A \in$ *proj-set* x A
    distproj x $A \in A$
    dist x (*distproj* x A) $=$ *infdist* x A
⟨proof⟩

lemma *proj-set-nonempty-of-proper*:
  assumes *proper* A $A \neq \{\}$
  shows *proj-set* x $A \neq \{\}$
⟨proof⟩

lemma *distproj-self* [*simp*]:
  assumes $x \in A$
  shows *proj-set* x $A = \{x\}$
    distproj x $A = x$
⟨proof⟩

lemma *distproj-closure* [*simp*]:
  assumes $x \in$ *closure* A
  shows *distproj* x $A = x$
⟨proof⟩

lemma *distproj-le*:
  assumes $y \in A$
  shows *dist* x (*distproj* x A) $\leq$ *dist* x y
⟨proof⟩

lemma *proj-set-dist-le*:
  assumes $y \in A$ $p \in$ *proj-set* x A
  shows *dist* x $p \leq$ *dist* x y
  ⟨proof⟩

1.3 Material on ereal and ennreal

We add the simp rules that we needed to make all computations become more or less automatic.

lemma *ereal-of-real-of-ereal-iff* [*simp*]:
$ereal(real\text{-}of\text{-}ereal\ x) = x \longleftrightarrow x \neq \infty \wedge x \neq -\infty$
$x = ereal(real\text{-}of\text{-}ereal\ x) \longleftrightarrow x \neq \infty \wedge x \neq -\infty$
⟨*proof*⟩

declare *ereal-inverse-eq-0* [*simp*]
declare *ereal-0-gt-inverse* [*simp*]
declare *ereal-inverse-le-0-iff* [*simp*]
declare *ereal-divide-eq-0-iff* [*simp*]
declare *ereal-mult-le-0-iff* [*simp*]
declare *ereal-zero-le-0-iff* [*simp*]
declare *ereal-mult-less-0-iff* [*simp*]
declare *ereal-zero-less-0-iff* [*simp*]
declare *ereal-uminus-eq-reorder* [*simp*]
declare *ereal-minus-le-iff* [*simp*]

lemma *ereal-inverse-noteq-minus-infinity* [*simp*]:
$1/(x::ereal) \neq -\infty$
⟨*proof*⟩

lemma *ereal-inverse-positive-iff-nonneg-not-infinity* [*simp*]:
$0 < 1/(x::ereal) \longleftrightarrow (x \geq 0 \wedge x \neq \infty)$
⟨*proof*⟩

lemma *ereal-inverse-negative-iff-nonpos-not-infinity'* [*simp*]:
$0 > inverse\ (x::ereal) \longleftrightarrow (x < 0 \wedge x \neq -\infty)$
⟨*proof*⟩

lemma *ereal-divide-pos-iff* [*simp*]:
$0 < x/(y::ereal) \longleftrightarrow (y \neq \infty \wedge y \neq -\infty) \wedge ((x > 0 \wedge y > 0) \vee (x < 0 \wedge y < 0) \vee (y = 0 \wedge x > 0))$
⟨*proof*⟩

lemma *ereal-divide-neg-iff* [*simp*]:
$0 > x/(y::ereal) \longleftrightarrow (y \neq \infty \wedge y \neq -\infty) \wedge ((x > 0 \wedge y < 0) \vee (x < 0 \wedge y > 0) \vee (y = 0 \wedge x < 0))$
⟨*proof*⟩

More additions to `mono_intros`.

lemma *ereal-leq-imp-neg-leq* [*mono-intros*]:
fixes $x\ y::ereal$
assumes $x \leq y$
shows $-y \leq -x$
⟨*proof*⟩

lemma *ereal-le-imp-neg-le* [*mono-intros*]:
fixes $x\ y::ereal$
assumes $x < y$
shows $-y < -x$

⟨*proof*⟩

declare *ereal-mult-left-mono* [*mono-intros*]
declare *ereal-mult-right-mono* [*mono-intros*]
declare *ereal-mult-strict-right-mono* [*mono-intros*]
declare *ereal-mult-strict-left-mono* [*mono-intros*]

Monotonicity of basic inclusions.

lemma *ennreal-mono'*:
mono ennreal
⟨*proof*⟩

lemma *enn2ereal-mono'*:
mono enn2ereal
⟨*proof*⟩

lemma *e2ennreal-mono'*:
mono e2ennreal
⟨*proof*⟩

lemma *enn2ereal-mono* [*mono-intros*]:
assumes $x \leq y$
shows *enn2ereal* $x \leq$ *enn2ereal* y
⟨*proof*⟩

lemma *ereal-mono*:
mono ereal
⟨*proof*⟩

lemma *ereal-strict-mono*:
strict-mono ereal
⟨*proof*⟩

lemma *ereal-mono2* [*mono-intros*]:
assumes $x \leq y$
shows *ereal* $x \leq$ *ereal* y
⟨*proof*⟩

lemma *ereal-strict-mono2* [*mono-intros*]:
assumes $x < y$
shows *ereal* $x <$ *ereal* y
⟨*proof*⟩

lemma *enn2ereal-a-minus-b-plus-b* [*mono-intros*]:
enn2ereal $a \leq$ *enn2ereal* $(a - b) +$ *enn2ereal* b
⟨*proof*⟩

The next lemma follows from the same assertion in ereals.

lemma *enn2ereal-strict-mono* [*mono-intros*]:

assumes $x < y$
shows *enn2ereal* $x <$ *enn2ereal* y
⟨*proof*⟩

declare *ennreal-mult-strict-left-mono* [*mono-intros*]
declare *ennreal-mult-strict-right-mono* [*mono-intros*]

lemma *ennreal-ge-0* [*mono-intros*]:
assumes $0 < x$
shows $0 <$ *ennreal* x
⟨*proof*⟩

The next lemma is true and useful in ereal. Note that variants such as $a + b \leq c + d$ implies $a - d \leq c - b$ are not true – take $a = c = \infty$ and $b = d = 0$...

lemma *ereal-minus-le-minus-plus* [*mono-intros*]:
fixes *a b c*::*ereal*
assumes $a \leq b + c$
shows $-b \leq -a + c$
⟨*proof*⟩

lemma *tendsto-ennreal-0* [*tendsto-intros*]:
assumes $(u \longrightarrow 0)\ F$
shows $((\lambda n.\ ennreal(u\ n)) \longrightarrow 0)\ F$
⟨*proof*⟩

lemma *tendsto-ennreal-1* [*tendsto-intros*]:
assumes $(u \longrightarrow 1)\ F$
shows $((\lambda n.\ ennreal(u\ n)) \longrightarrow 1)\ F$
⟨*proof*⟩

1.4 Miscellaneous

lemma *lim-ceiling-over-n* [*tendsto-intros*]:
assumes $(\lambda n.\ u\ n/n) \longlonglongrightarrow l$
shows $(\lambda n.\ ceiling(u\ n)/n) \longlonglongrightarrow l$
⟨*proof*⟩

1.4.1 Liminfs and Limsups

More facts on liminfs and limsups

lemma *Limsup-obtain'*:
fixes *u*::*'a* ⇒ *'b*::*complete-linorder*
assumes *Limsup F u* $> c$ *eventually P F*
shows $\exists n.\ P\ n \wedge u\ n > c$
⟨*proof*⟩

lemma *limsup-obtain*:

fixes $u::nat \Rightarrow 'a :: complete\text{-}linorder$
assumes $limsup\ u > c$
shows $\exists\, n \geq N.\ u\ n > c$
⟨*proof*⟩

lemma *Liminf-obtain'*:
fixes $u::'a \Rightarrow 'b::complete\text{-}linorder$
assumes $Liminf\ F\ u < c$ $eventually\ P\ F$
shows $\exists\, n.\ P\ n \wedge u\ n < c$
⟨*proof*⟩

lemma *liminf-obtain*:
fixes $u::nat \Rightarrow 'a :: complete\text{-}linorder$
assumes $liminf\ u < c$
shows $\exists\, n \geq N.\ u\ n < c$
⟨*proof*⟩

The Liminf of a minimum is the minimum of the Liminfs.

lemma *Liminf-min-eq-min-Liminf*:
fixes $u\ v::nat \Rightarrow 'a::complete\text{-}linorder$
shows $Liminf\ F\ (\lambda n.\ min\ (u\ n)\ (v\ n)) = min\ (Liminf\ F\ u)\ (Liminf\ F\ v)$
⟨*proof*⟩

The Limsup of a maximum is the maximum of the Limsups.

lemma *Limsup-max-eq-max-Limsup*:
fixes $u::'a \Rightarrow 'b::complete\text{-}linorder$
shows $Limsup\ F\ (\lambda n.\ max\ (u\ n)\ (v\ n)) = max\ (Limsup\ F\ u)\ (Limsup\ F\ v)$
⟨*proof*⟩

1.4.2 Bounding the cardinality of a finite set

A variation with real bounds.

lemma *finite-finite-subset-caract'*:
fixes $C::real$
assumes $\bigwedge G.\ G \subseteq F \Longrightarrow finite\ G \Longrightarrow card\ G \leq C$
shows $finite\ F \wedge card\ F \leq C$
⟨*proof*⟩

To show that a set has cardinality at most one, it suffices to show that any two of its elements coincide.

lemma *finite-at-most-singleton*:
assumes $\bigwedge x\ y.\ x \in F \Longrightarrow y \in F \Longrightarrow x = y$
shows $finite\ F \wedge card\ F \leq 1$
⟨*proof*⟩

Bounded sets of integers are finite.

lemma *finite-real-int-interval* [*simp*]:
$finite\ (range\ real\text{-}of\text{-}int \cap \{a..b\})$

⟨*proof*⟩

Well separated sets of real numbers are finite, with controlled cardinality.

lemma *separated-in-real-card-bound*:
  assumes $T \subseteq \{a..(b::real)\}$ $d > 0$ $\bigwedge x\ y.\ x \in T \Longrightarrow y \in T \Longrightarrow y > x \Longrightarrow y \geq x + d$
  shows *finite T* $card\ T \leq nat\ (floor\ ((b-a)/d) + 1)$
⟨*proof*⟩

1.5 Manipulating finite ordered sets

We will need below to construct finite sets of real numbers with good properties expressed in terms of consecutive elements of the set. We introduce tools to manipulate such sets, expressing in particular the next and the previous element of the set and controlling how they evolve when one inserts a new element in the set. It works in fact in any linorder, and could also prove useful to construct sets of integer numbers.

Manipulating the next and previous elements work well, except at the top (respectively bottom). In our constructions, these will be fixed and called b and a.

Notations for the next and the previous elements.

definition $next\text{-}in::'a\ set \Rightarrow 'a \Rightarrow ('a::linorder)$
  where $next\text{-}in\ A\ u = Min\ (A \cap \{u<..\})$

definition $prev\text{-}in::'a\ set \Rightarrow 'a \Rightarrow ('a::linorder)$
  where $prev\text{-}in\ A\ u = Max\ (A \cap \{..<u\})$

context
  fixes $A::('a::linorder)\ set$ **and** $a\ b::'a$
  assumes A: *finite A* $A \subseteq \{a..b\}$ $a \in A$ $b \in A$ $a < b$
begin

Basic properties of the next element, when one starts from an element different from top.

lemma *next-in-basics*:
  assumes $u \in \{a..<b\}$
  shows $next\text{-}in\ A\ u \in A$
    $next\text{-}in\ A\ u > u$
    $A \cap \{u<..<next\text{-}in\ A\ u\} = \{\}$
⟨*proof*⟩

lemma *next-inI*:
  assumes $u \in \{a..<b\}$
    $v \in A$
    $v > u$
    $\{u<..<v\} \cap A = \{\}$

shows *next-in A u* = *v*
⟨*proof*⟩

Basic properties of the previous element, when one starts from an element different from bottom.

lemma *prev-in-basics*:
assumes $u \in \{a<..b\}$
shows $prev\text{-}in\ A\ u \in A$
$prev\text{-}in\ A\ u < u$
$A \cap \{prev\text{-}in\ A\ u<..<u\} = \{\}$
⟨*proof*⟩

lemma *prev-inI*:
assumes $u \in \{a<..b\}$
$v \in A$
$v < u$
$\{v<..<u\} \cap A = \{\}$
shows $prev\text{-}in\ A\ u = v$
⟨*proof*⟩

The interval $[a, b]$ is covered by the intervals between the consecutive elements of A.

lemma *intervals-decomposition*:
$(\bigcup\ U \in \{\{u..next\text{-}in\ A\ u\} \mid u.\ u \in A - \{b\}\}.\ U) = \{a..b\}$
⟨*proof*⟩
end

If one inserts an additional element, then next and previous elements are not modified, except at the location of the insertion.

lemma *next-in-insert*:
assumes A: $finite\ A\ A \subseteq \{a..b\}\ a \in A\ b \in A\ a < b$
and $x \in \{a..b\} - A$
shows $\bigwedge u.\ u \in A - \{b, prev\text{-}in\ A\ x\} \Longrightarrow next\text{-}in\ (insert\ x\ A)\ u = next\text{-}in\ A\ u$
$next\text{-}in\ (insert\ x\ A)\ x = next\text{-}in\ A\ x$
$next\text{-}in\ (insert\ x\ A)\ (prev\text{-}in\ A\ x) = x$
⟨*proof*⟩

If consecutive elements are enough separated, this implies a simple bound on the cardinality of the set.

lemma *separated-in-real-card-bound2*:
fixes A::*real set*
assumes A: $finite\ A\ A \subseteq \{a..b\}\ a \in A\ b \in A\ a < b$
and B: $\bigwedge u.\ u \in A - \{b\} \Longrightarrow next\text{-}in\ A\ u \geq u + d\ d > 0$
shows $card\ A \leq nat\ (floor\ ((b-a)/d) + 1)$
⟨*proof*⟩

1.6 Well-orders

In this subsection, we give additional lemmas on well-orders or cardinals or whatever, that would well belong to the library, and will be needed below.

lemma (**in** *wo-rel*) *max2-underS* [*simp*]:
  assumes $x \in$ *underS* z $y \in$ *underS* z
  shows *max2* x $y \in$ *underS* z
⟨*proof*⟩

lemma (**in** *wo-rel*) *max2-underS′* [*simp*]:
  assumes $x \in$ *underS* y
  shows *max2* x $y = y$ *max2* y $x = y$
⟨*proof*⟩

lemma (**in** *wo-rel*) *max2-xx* [*simp*]:
  max2 x $x = x$
⟨*proof*⟩

declare *underS-notIn* [*simp*]

The abbrevation $=o$ is used both in `Set_Algebras` and Cardinals. We disable the one from `Set_Algebras`.

no-notation *elt-set-eq* (**infix** ‹=*o*› *50*)

lemma *regularCard-ordIso*:
  assumes *Card-order* r *regularCard* r $s =o$ r
  shows *regularCard* s
⟨*proof*⟩

lemma *AboveS-not-empty-in-regularCard*:
  assumes $|S| <o$ r $S \subseteq$ *Field* r
  assumes r: *Card-order* r *regularCard* r $\neg$*finite* (*Field* r)
  shows *AboveS* r $S \neq \{\}$
⟨*proof*⟩

lemma *AboveS-not-empty-in-regularCard′*:
  assumes $|S| <o$ r $f`S \subseteq$ *Field* r $T \subseteq S$
  assumes r: *Card-order* r *regularCard* r $\neg$*finite* (*Field* r)
  shows *AboveS* r $(f`T) \neq \{\}$
⟨*proof*⟩

lemma *Well-order-extend*:
assumes *WELL*: *well-order-on* A r **and** *SUB*: $A \subseteq B$
shows $\exists r'.$ *well-order-on* B $r' \wedge r \subseteq r'$
⟨*proof*⟩

The next lemma shows that, if the range of a function is endowed with a wellorder, then one can pull back this wellorder by the function, and then extend it in the fibers of the function in order to keep the wellorder property.

The proof is done by taking an arbitrary family of wellorders on each of the fibers, and using the lexicographic order: one has $x < y$ if $fx < fy$, or if $fx = fy$ and, in the corresponding fiber of f, one has $x < y$.

To formalize it, it is however more efficient to use one single wellorder, and restrict it to each fiber.

lemma *Well-order-pullback*:
  assumes *Well-order r*
  shows $\exists\, s.$ *Well-order s* $\wedge$ *Field s* $=$ *UNIV* $\wedge$ $(\forall\, x\, y.\ (f\, x,\, f\, y) \in (r{-}Id) \longrightarrow (x,\, y) \in s)$
⟨*proof*⟩

end

2 The exponential on extended real numbers.

theory *Eexp-Eln*
  imports *Library-Complements*
begin

To define the distance on the Gromov completion of hyperbolic spaces, we need to use the exponential on extended real numbers. We can not use the symbol `exp`, as this symbol is already used in Banach algebras, so we use `ennexp` instead. We prove its basic properties (together with properties of the logarithm) here. We also use it to define the square root on ennreal. Finally, we also define versions from ereal to ereal.

function *ennexp*::*ereal* ⇒ *ennreal* **where**
ennexp (*ereal r*) = *ennreal* (*exp r*)
| *ennexp* (∞) = ∞
| *ennexp* (−∞) = *0*
⟨*proof*⟩
termination ⟨*proof*⟩

lemma *ennexp-0* [*simp*]:
  ennexp 0 = *1*
⟨*proof*⟩

function *eln*::*ennreal* ⇒ *ereal* **where**
eln (*ennreal r*) = (*if r* ≤ *0 then* −∞ *else ereal* (*ln r*))
| *eln* (∞) = ∞
⟨*proof*⟩
termination ⟨*proof*⟩

lemma *eln-simps* [*simp*]:
  eln 0 = −∞
  eln 1 = *0*
  eln top = ∞
⟨*proof*⟩

lemma *eln-real-pos*:
  assumes $r > 0$
  shows *eln* (*ennreal* r) = *ereal* (*ln* r)
⟨*proof*⟩

lemma *eln-ennexp* [*simp*]:
  eln (*ennexp* x) = x
⟨*proof*⟩

lemma *ennexp-eln* [*simp*]:
  ennexp (*eln* x) = x
⟨*proof*⟩

lemma *ennexp-strict-mono*:
  strict-mono ennexp
⟨*proof*⟩

lemma *ennexp-mono*:
  mono ennexp
⟨*proof*⟩

lemma *ennexp-strict-mono2* [*mono-intros*]:
  assumes $x < y$
  shows *ennexp* x < *ennexp* y
⟨*proof*⟩

lemma *ennexp-mono2* [*mono-intros*]:
  assumes $x \leq y$
  shows *ennexp* $x \leq$ *ennexp* y
⟨*proof*⟩

lemma *ennexp-le1* [*simp*]:
  ennexp $x \leq 1 \longleftrightarrow x \leq 0$
⟨*proof*⟩

lemma *ennexp-ge1* [*simp*]:
  ennexp $x \geq 1 \longleftrightarrow x \geq 0$
⟨*proof*⟩

lemma *eln-strict-mono*:
  strict-mono eln
⟨*proof*⟩

lemma *eln-mono*:
  mono eln
⟨*proof*⟩

lemma *eln-strict-mono2* [*mono-intros*]:

assumes $x < y$
shows *eln* $x <$ *eln* y
⟨*proof*⟩

lemma *eln-mono2* [*mono-intros*]:
assumes $x \leq y$
shows *eln* $x \leq$ *eln* y
⟨*proof*⟩

lemma *eln-le0* [*simp*]:
eln $x \leq 0 \longleftrightarrow x \leq 1$
⟨*proof*⟩

lemma *eln-ge0* [*simp*]:
eln $x \geq 0 \longleftrightarrow x \geq 1$
⟨*proof*⟩

lemma *bij-ennexp*:
bij ennexp
⟨*proof*⟩

lemma *bij-eln*:
bij eln
⟨*proof*⟩

lemma *ennexp-continuous*:
continuous-on UNIV ennexp
⟨*proof*⟩

lemma *ennexp-tendsto* [*tendsto-intros*]:
assumes $((\lambda n.\ u\ n) \longrightarrow l)\ F$
shows $((\lambda n.\ ennexp(u\ n)) \longrightarrow ennexp\ l)\ F$
⟨*proof*⟩

lemma *eln-continuous*:
continuous-on UNIV eln
⟨*proof*⟩

lemma *eln-tendsto* [*tendsto-intros*]:
assumes $((\lambda n.\ u\ n) \longrightarrow l)\ F$
shows $((\lambda n.\ eln(u\ n)) \longrightarrow eln\ l)\ F$
⟨*proof*⟩

lemma *ennexp-special-values* [*simp*]:
ennexp $x = 0 \longleftrightarrow x = -\infty$
ennexp $x = 1 \longleftrightarrow x = 0$
ennexp $x = \infty \longleftrightarrow x = \infty$
ennexp $x = top \longleftrightarrow x = \infty$
⟨*proof*⟩

lemma *eln-special-values* [*simp*]:
  eln $x = -\infty \longleftrightarrow x = 0$
  eln $x = 0 \longleftrightarrow x = 1$
  eln $x = \infty \longleftrightarrow x = \infty$
⟨*proof*⟩

lemma *ennexp-add-mult*:
  assumes $\neg((a = \infty \wedge b = -\infty) \vee (a = -\infty \wedge b = \infty))$
  shows *ennexp*$(a+b)$ = *ennexp* a ∗ *ennexp* b
⟨*proof*⟩

lemma *eln-mult-add*:
  assumes $\neg((a = \infty \wedge b = 0) \vee (a = 0 \wedge b = \infty))$
  shows *eln*$(a * b)$ = *eln* a + *eln* b
⟨*proof*⟩

We can also define the square root on ennreal using the above exponential.

definition *ennsqrt*::*ennreal* ⇒ *ennreal*
  where *ennsqrt* x = *ennexp*(*eln* $x/2$)

lemma *ennsqrt-square* [*simp*]:
  (*ennsqrt* x) ∗ (*ennsqrt* x) = x
⟨*proof*⟩

lemma *ennsqrt-simps* [*simp*]:
  ennsqrt $0 = 0$
  ennsqrt $1 = 1$
  ennsqrt $\infty = \infty$
  ennsqrt top = *top*
⟨*proof*⟩

lemma *ennsqrt-mult*:
  ennsqrt$(a * b)$ = *ennsqrt* a ∗ *ennsqrt* b
⟨*proof*⟩

lemma *ennsqrt-square2* [*simp*]:
  ennsqrt $(x * x) = x$
  ⟨*proof*⟩

lemma *ennsqrt-eq-iff-square*:
  ennsqrt $x = y \longleftrightarrow x = y * y$
⟨*proof*⟩

lemma *ennsqrt-bij*:
  bij ennsqrt
⟨*proof*⟩

lemma *ennsqrt-strict-mono*:

strict-mono ennsqrt
⟨*proof*⟩

lemma *ennsqrt-mono*:
mono ennsqrt
⟨*proof*⟩

lemma *ennsqrt-mono2* [*mono-intros*]:
assumes $x \leq y$
shows *ennsqrt* $x \leq$ *ennsqrt* y
⟨*proof*⟩

lemma *ennsqrt-continuous*:
continuous-on UNIV ennsqrt
⟨*proof*⟩

lemma *ennsqrt-tendsto* [*tendsto-intros*]:
assumes $((\lambda n.\ u\ n) \longrightarrow l)\ F$
shows $((\lambda n.\ ennsqrt(u\ n)) \longrightarrow ennsqrt\ l)\ F$
⟨*proof*⟩

lemma *ennsqrt-ennreal-ennreal-sqrt* [*simp*]:
assumes $t \geq (0::real)$
shows *ennsqrt* (*ennreal* t) = *ennreal* (*sqrt* t)
⟨*proof*⟩

lemma *ennreal-sqrt2*:
ennreal (*sqrt 2*) = *ennsqrt 2*
⟨*proof*⟩

lemma *ennsqrt-4* [*simp*]:
ennsqrt 4 = *2*
⟨*proof*⟩

lemma *ennsqrt-le* [*simp*]:
ennsqrt $x \leq$ *ennsqrt* $y \longleftrightarrow x \leq y$
⟨*proof*⟩

We can also define the square root on ereal using the square root on ennreal, and 0 for negative numbers.

definition *esqrt*::*ereal* ⇒ *ereal*
where *esqrt* x = *enn2ereal*(*ennsqrt* (*e2ennreal* x))

lemma *esqrt-square* [*simp*]:
assumes $x \geq 0$
shows (*esqrt* x) ∗ (*esqrt* x) = x
⟨*proof*⟩

lemma *esqrt-of-neg* [*simp*]:

assumes $x \leq 0$
shows $esqrt\ x = 0$
⟨*proof*⟩

lemma *esqrt-nonneg* [*simp*]:
$esqrt\ x \geq 0$
⟨*proof*⟩

lemma *esqrt-eq-iff-square* [*simp*]:
assumes $x \geq 0$ $y \geq 0$
shows $esqrt\ x = y \longleftrightarrow x = y * y$
⟨*proof*⟩

lemma *esqrt-simps* [*simp*]:
$esqrt\ 0 = 0$
$esqrt\ 1 = 1$
$esqrt\ \infty = \infty$
$esqrt\ top = top$
$esqrt\ (-\infty) = 0$
⟨*proof*⟩

lemma *esqrt-mult*:
assumes $a \geq 0$
shows $esqrt(a * b) = esqrt\ a * esqrt\ b$
⟨*proof*⟩

lemma *esqrt-square2* [*simp*]:
$esqrt(x * x) = abs(x)$
⟨*proof*⟩

lemma *esqrt-mono*:
mono esqrt
⟨*proof*⟩

lemma *esqrt-mono2* [*mono-intros*]:
assumes $x \leq y$
shows $esqrt\ x \leq esqrt\ y$
⟨*proof*⟩

lemma *esqrt-continuous*:
continuous-on UNIV esqrt
⟨*proof*⟩

lemma *esqrt-tendsto* [*tendsto-intros*]:
assumes $((\lambda n.\ u\ n) \longrightarrow l)\ F$
shows $((\lambda n.\ esqrt(u\ n)) \longrightarrow esqrt\ l)\ F$
⟨*proof*⟩

lemma *esqrt-ereal-ereal-sqrt* [*simp*]:

assumes $t \geq (0::real)$
shows $esqrt\ (ereal\ t) = ereal\ (sqrt\ t)$
⟨*proof*⟩

lemma *ereal-sqrt2*:
$ereal\ (sqrt\ 2) = esqrt\ 2$
⟨*proof*⟩

lemma *esqrt-4* [*simp*]:
$esqrt\ 4 = 2$
⟨*proof*⟩

lemma *esqrt-le* [*simp*]:
$esqrt\ x \leq esqrt\ y \longleftrightarrow (x \leq 0 \lor x \leq y)$
⟨*proof*⟩

Finally, we define eexp, as the composition of ennexp and the injection of ennreal in ereal.

definition *eexp*::*ereal* ⇒ *ereal* **where**
$eexp\ x = enn2ereal\ (ennexp\ x)$

lemma *eexp-special-values* [*simp*]:
$eexp\ 0 = 1$
$eexp\ (\infty) = \infty$
$eexp(-\infty) = 0$
⟨*proof*⟩

lemma *eexp-strict-mono*:
strict-mono eexp
⟨*proof*⟩

lemma *eexp-mono*:
mono eexp
⟨*proof*⟩

lemma *eexp-strict-mono2* [*mono-intros*]:
assumes $x < y$
shows $eexp\ x < eexp\ y$
⟨*proof*⟩

lemma *eexp-mono2* [*mono-intros*]:
assumes $x \leq y$
shows $eexp\ x \leq eexp\ y$
⟨*proof*⟩

lemma *eexp-le-eexp-iff-le*:
$eexp\ x \leq eexp\ y \longleftrightarrow x \leq y$
⟨*proof*⟩

lemma *eexp-lt-eexp-iff-lt*:
$eexp\ x < eexp\ y \longleftrightarrow x < y$
⟨*proof*⟩

lemma *eexp-special-values-iff* [*simp*]:
$eexp\ x = 0 \longleftrightarrow x = -\infty$
$eexp\ x = 1 \longleftrightarrow x = 0$
$eexp\ x = \infty \longleftrightarrow x = \infty$
$eexp\ x = top \longleftrightarrow x = \infty$
⟨*proof*⟩

lemma *eexp-ineq-iff* [*simp*]:
$eexp\ x \leq 1 \longleftrightarrow x \leq 0$
$eexp\ x \geq 1 \longleftrightarrow x \geq 0$
$eexp\ x > 1 \longleftrightarrow x > 0$
$eexp\ x < 1 \longleftrightarrow x < 0$
$eexp\ x \geq 0$
$eexp\ x > 0 \longleftrightarrow x \neq -\infty$
$eexp\ x < \infty \longleftrightarrow x \neq \infty$
⟨*proof*⟩

lemma *eexp-ineq* [*mono-intros*]:
$x \leq 0 \Longrightarrow eexp\ x \leq 1$
$x < 0 \Longrightarrow eexp\ x < 1$
$x \geq 0 \Longrightarrow eexp\ x \geq 1$
$x > 0 \Longrightarrow eexp\ x > 1$
$eexp\ x \geq 0$
$x > -\infty \Longrightarrow eexp\ x > 0$
$x < \infty \Longrightarrow eexp\ x < \infty$
⟨*proof*⟩

lemma *eexp-continuous*:
continuous-on UNIV eexp
⟨*proof*⟩

lemma *eexp-tendsto'* [*simp*]:
$((\lambda n.\ eexp(u\ n)) \longrightarrow eexp\ l)\ F \longleftrightarrow ((\lambda n.\ u\ n) \longrightarrow l)\ F$
⟨*proof*⟩

lemma *eexp-tendsto* [*tendsto-intros*]:
assumes $((\lambda n.\ u\ n) \longrightarrow l)\ F$
shows $((\lambda n.\ eexp(u\ n)) \longrightarrow eexp\ l)\ F$
⟨*proof*⟩

lemma *eexp-add-mult*:
assumes $\neg((a = \infty \wedge b = -\infty) \vee (a = -\infty \wedge b = \infty))$
shows $eexp(a+b) = eexp\ a * eexp\ b$
⟨*proof*⟩

lemma *eexp-ereal* [*simp*]:
  eexp(*ereal x*) = *ereal*(*exp x*)
⟨*proof*⟩

end

3 Hausdorff distance

theory *Hausdorff-Distance*
  imports *Library-Complements*
begin

3.1 Preliminaries

3.2 Hausdorff distance

The Hausdorff distance between two subsets of a metric space is the minimal M such that each set is included in the M-neighborhood of the other. For nonempty bounded sets, it satisfies the triangular inequality, it is symmetric, but it vanishes on sets that have the same closure. In particular, it defines a distance on closed bounded nonempty sets. We establish all these properties below.

definition *hausdorff-distance*::(*'a*::*metric-space*) *set* ⇒ *'a set* ⇒ *real*
  where *hausdorff-distance A B* = (*if A* = {} ∨ *B* = {} ∨ (¬(*bounded A*)) ∨ (¬(*bounded B*)) *then 0*
            else max (*SUP x*∈*A*. *infdist x B*) (*SUP x*∈*B*. *infdist x A*))

lemma *hausdorff-distance-self* [*simp*]:
  hausdorff-distance A A = *0*
⟨*proof*⟩

lemma *hausdorff-distance-sym*:
  hausdorff-distance A B = *hausdorff-distance B A*
⟨*proof*⟩

lemma *hausdorff-distance-points* [*simp*]:
  hausdorff-distance {*x*} {*y*} = *dist x y*
⟨*proof*⟩

The Hausdorff distance is expressed in terms of a supremum. To use it, one needs again and again to show that this is the supremum of a set which is bounded from above.

lemma *bdd-above-infdist-aux*:
  assumes *bounded A bounded B*
  shows *bdd-above* ((λ*x*. *infdist x B*)‘*A*)

⟨*proof*⟩

lemma *hausdorff-distance-nonneg* [*simp*, *mono-intros*]:
hausdorff-distance A B $\geq$ *0*
⟨*proof*⟩

lemma *hausdorff-distanceI*:
assumes $\bigwedge x.\ x \in A \Longrightarrow$ *infdist x B* $\leq$ *D*
$\bigwedge x.\ x \in B \Longrightarrow$ *infdist x A* $\leq$ *D*
D $\geq$ *0*
shows *hausdorff-distance A B* $\leq$ *D*
⟨*proof*⟩

lemma *hausdorff-distanceI2*:
assumes $\bigwedge x.\ x \in A \Longrightarrow \exists\, y \in B.$ *dist x y* $\leq$ *D*
$\bigwedge x.\ x \in B \Longrightarrow \exists\, y \in A.$ *dist x y* $\leq$ *D*
D $\geq$ *0*
shows *hausdorff-distance A B* $\leq$ *D*
⟨*proof*⟩

lemma *infdist-le-hausdorff-distance* [*mono-intros*]:
assumes $x \in A$ *bounded A bounded B*
shows *infdist x B* $\leq$ *hausdorff-distance A B*
⟨*proof*⟩

lemma *hausdorff-distance-infdist-triangle* [*mono-intros*]:
assumes $B \neq \{\}$ *bounded B bounded C*
shows *infdist x C* $\leq$ *infdist x B* + *hausdorff-distance B C*
⟨*proof*⟩

lemma *hausdorff-distance-triangle* [*mono-intros*]:
assumes $B \neq \{\}$ *bounded B*
shows *hausdorff-distance A C* $\leq$ *hausdorff-distance A B* + *hausdorff-distance B C*
⟨*proof*⟩

lemma *hausdorff-distance-subset*:
assumes $A \subseteq B$ $A \neq \{\}$ *bounded B*
shows *hausdorff-distance A B* = (*SUP* $x \in B$. *infdist x A*)
⟨*proof*⟩

lemma *hausdorff-distance-closure* [*simp*]:
hausdorff-distance A (*closure A*) = *0*
⟨*proof*⟩

lemma *hausdorff-distance-closures* [*simp*]:
hausdorff-distance (*closure A*) (*closure B*) = *hausdorff-distance A B*
⟨*proof*⟩

lemma *hausdorff-distance-zero*:
  assumes $A \neq \{\}$ *bounded* A $B \neq \{\}$ *bounded* B
  shows *hausdorff-distance* A $B = 0 \longleftrightarrow$ *closure* $A =$ *closure* B
⟨*proof*⟩

lemma *hausdorff-distance-vimage*:
  assumes $\bigwedge x.\ x \in A \Longrightarrow$ *dist* $(f\ x)\ (g\ x) \leq C$
    $C \geq 0$
  shows *hausdorff-distance* $(f`A)$ $(g`A) \leq C$
⟨*proof*⟩

lemma *hausdorff-distance-union* [*mono-intros*]:
  assumes $A \neq \{\}$ $B \neq \{\}$ $C \neq \{\}$ $D \neq \{\}$
  shows *hausdorff-distance* $(A \cup B)$ $(C \cup D) \leq$ *max* (*hausdorff-distance* A C) (*hausdorff-distance* B D)
⟨*proof*⟩

end

4 Isometries

theory *Isometries*
  imports *Library-Complements Hausdorff-Distance*
begin

Isometries, i.e., functions that preserve distances, show up very often in mathematics. We introduce a dedicated definition, and show its basic properties.

definition *isometry-on*::('*a*::*metric-space*) *set* ⇒ ('*a* ⇒ ('*b*::*metric-space*)) ⇒ *bool*
  where *isometry-on* X $f = (\forall x \in X.\ \forall y \in X.$ *dist* $(f\ x)\ (f\ y) =$ *dist* $x\ y)$

definition *isometry* :: ('*a*::*metric-space* ⇒ '*b*::*metric-space*) ⇒ *bool*
  where *isometry* $f \equiv$ *isometry-on* *UNIV* $f \wedge$ *range* $f =$ *UNIV*

lemma *isometry-on-subset*:
  assumes *isometry-on* X f
    $Y \subseteq X$
  shows *isometry-on* Y f
⟨*proof*⟩

lemma *isometry-onI* [*intro?*]:
  assumes $\bigwedge x\ y.\ x \in X \Longrightarrow y \in X \Longrightarrow$ *dist* $(f\ x)\ (f\ y) =$ *dist* $x\ y$
  shows *isometry-on* X f
⟨*proof*⟩

lemma *isometry-onD*:
  assumes *isometry-on* X f
    $x \in X$ $y \in X$

shows *dist* (*f x*) (*f y*) = *dist x y*
⟨*proof*⟩

lemma *isometryI* [*intro?*]:
assumes $\bigwedge x\ y.$ *dist* (*f x*) (*f y*) = *dist x y*
range f = *UNIV*
shows *isometry f*
⟨*proof*⟩

lemma
assumes *isometry-on X f*
shows *isometry-on-lipschitz*: *1*−*lipschitz-on X f*
and *isometry-on-uniformly-continuous*: *uniformly-continuous-on X f*
and *isometry-on-continuous*: *continuous-on X f*
⟨*proof*⟩

lemma *isometryD*:
assumes *isometry f*
shows *isometry-on UNIV f*
dist (*f x*) (*f y*) = *dist x y*
range f = *UNIV*
1−*lipschitz-on UNIV f*
uniformly-continuous-on UNIV f
continuous-on UNIV f
⟨*proof*⟩

lemma *isometry-on-injective*:
assumes *isometry-on X f*
shows *inj-on f X*
⟨*proof*⟩

lemma *isometry-on-compose*:
assumes *isometry-on X f*
isometry-on (*f'X*) *g*
shows *isometry-on X* (λ*x*. *g*(*f x*))
⟨*proof*⟩

lemma *isometry-on-cong*:
assumes *isometry-on X f*
$\bigwedge x.\ x \in X \Longrightarrow g\ x = f\ x$
shows *isometry-on X g*
⟨*proof*⟩

lemma *isometry-on-inverse*:
assumes *isometry-on X f*
shows *isometry-on* (*f'X*) (*inv-into X f*)
$\bigwedge x.\ x \in X \Longrightarrow$ (*inv-into X f*) (*f x*) = *x*
$\bigwedge y.\ y \in f`X \Longrightarrow$ *f* (*inv-into X f y*) = *y*
bij-betw f X (*f'X*)

⟨*proof*⟩

lemma *isometry-inverse*:
  assumes *isometry f*
  shows *isometry* (*inv f*)
    bij f
⟨*proof*⟩

lemma *isometry-on-homeomorphism*:
  assumes *isometry-on X f*
  shows *homeomorphism X* (*f'X*) *f* (*inv-into X f*)
    homeomorphism-on X f
    X homeomorphic f'X
⟨*proof*⟩

lemma *isometry-homeomorphism*:
  fixes *f*::(*'a*::*metric-space*) ⇒ (*'b*::*metric-space*)
  assumes *isometry f*
  shows *homeomorphism UNIV UNIV f* (*inv f*)
    (*UNIV*::*'a set*) *homeomorphic* (*UNIV*::*'b set*)
⟨*proof*⟩

lemma *isometry-on-closure*:
  assumes *isometry-on X f*
    continuous-on (*closure X*) *f*
  shows *isometry-on* (*closure X*) *f*
⟨*proof*⟩

lemma *isometry-extend-closure*:
  fixes *f*::(*'a*::*metric-space*) ⇒ (*'b*::*complete-space*)
  assumes *isometry-on X f*
  shows ∃ *g*. *isometry-on* (*closure X*) *g* ∧ (∀ *x*∈*X*. *g x* = *f x*)
⟨*proof*⟩

lemma *isometry-on-complete-image*:
  assumes *isometry-on X f*
    complete X
  shows *complete* (*f'X*)
⟨*proof*⟩

lemma *isometry-on-id* [*simp*]:
  isometry-on A (λ*x*. *x*)
  isometry-on A id
⟨*proof*⟩

lemma *isometry-on-add* [*simp*]:
  isometry-on A (λ*x*. *x* + (*t*::*'a*::*real-normed-vector*))
⟨*proof*⟩

lemma *isometry-on-minus* [*simp*]:
isometry-on A (λ(x::$'a$::*real-normed-vector*). $-x$)
⟨*proof*⟩

lemma *isometry-on-diff* [*simp*]:
isometry-on A (λx. (t::$'a$::*real-normed-vector*) $-$ x)
⟨*proof*⟩

lemma *isometry-preserves-bounded*:
assumes *isometry-on* X f
$A \subseteq X$
shows *bounded* (f‘A) $\longleftrightarrow$ *bounded* A
⟨*proof*⟩

lemma *isometry-preserves-infdist*:
infdist (f x) (f‘A) = *infdist* x A
if *isometry-on* X f $A \subseteq X$ $x \in X$
⟨*proof*⟩

lemma *isometry-preserves-hausdorff-distance*:
hausdorff-distance (f‘A) (f‘B) = *hausdorff-distance* A B
if *isometry-on* X f $A \subseteq X$ $B \subseteq X$
⟨*proof*⟩

lemma *isometry-on-UNIV-iterates*:
fixes f::($'a$::*metric-space*) $\Rightarrow$ $'a$
assumes *isometry-on* *UNIV* f
shows *isometry-on* *UNIV* (f^^n)
⟨*proof*⟩

lemma *isometry-iterates*:
fixes f::($'a$::*metric-space*) $\Rightarrow$ $'a$
assumes *isometry* f
shows *isometry* (f^^n)
⟨*proof*⟩

5 Geodesic spaces

A geodesic space is a metric space in which any pair of points can be joined by a geodesic segment, i.e., an isometrically embedded copy of a segment in the real line. Most spaces in geometry are geodesic. We introduce in this section the corresponding class of metric spaces. First, we study properties of general geodesic segments in metric spaces.

5.1 Geodesic segments in general metric spaces

definition *geodesic-segment-between*::($'a$::*metric-space*) *set* $\Rightarrow$ $'a$ $\Rightarrow$ $'a$ $\Rightarrow$ *bool*

where *geodesic-segment-between G x y* = (∃ *g*::(*real* ⇒ ′*a*). *g 0* = *x* ∧ *g* (*dist x y*) = *y* ∧ *isometry-on* {*0..dist x y*} *g* ∧ *G* = *g*‘{*0..dist x y*})

definition *geodesic-segment*::(′*a*::*metric-space*) *set* ⇒ *bool*
where *geodesic-segment G* = (∃ *x y*. *geodesic-segment-between G x y*)

We also introduce the parametrization of a geodesic segment. It is convenient to use the following definition, which guarantees that the point is on G even without checking that G is a geodesic segment or that the parameter is in the reasonable range: this shortens some arguments below.

definition *geodesic-segment-param*::(′*a*::*metric-space*) *set* ⇒ ′*a* ⇒ *real* ⇒ ′*a*
where *geodesic-segment-param G x t* = (*if* ∃ *w*. *w* ∈ *G* ∧ *dist x w* = *t then SOME w*. *w* ∈ *G* ∧ *dist x w* = *t else SOME w*. *w* ∈ *G*)

lemma *geodesic-segment-betweenI*:
assumes *g 0* = *x g* (*dist x y*) = *y isometry-on* {*0..dist x y*} *g G* = *g*‘{*0..dist x y*}
shows *geodesic-segment-between G x y*
⟨*proof*⟩

lemma *geodesic-segmentI* [*intro*, *simp*]:
assumes *geodesic-segment-between G x y*
shows *geodesic-segment G*
⟨*proof*⟩

lemma *geodesic-segmentI2* [*intro*]:
assumes *isometry-on* {*a..b*} *g a* ≤ (*b*::*real*)
shows *geodesic-segment-between* (*g*‘{*a..b*}) (*g a*) (*g b*)
geodesic-segment (*g*‘{*a..b*})
⟨*proof*⟩

lemma *geodesic-segmentD*:
assumes *geodesic-segment-between G x y*
shows ∃ *g*::(*real* ⇒ -). (*g t* = *x* ∧ *g* (*t* + *dist x y*) = *y* ∧ *isometry-on* {*t..t+dist x y*} *g* ∧ *G* = *g*‘{*t..t+dist x y*})
⟨*proof*⟩

lemma *geodesic-segment-endpoints* [*simp*]:
assumes *geodesic-segment-between G x y*
shows *x* ∈ *G y* ∈ *G G* ≠ {}
⟨*proof*⟩

lemma *geodesic-segment-commute*:
assumes *geodesic-segment-between G x y*
shows *geodesic-segment-between G y x*
⟨*proof*⟩

lemma *geodesic-segment-dist*:
assumes *geodesic-segment-between G x y a* ∈ *G*

shows *dist x a* + *dist a y* = *dist x y*
⟨*proof*⟩

lemma *geodesic-segment-dist-unique*:
assumes *geodesic-segment-between G x y a* ∈ *G b* ∈ *G dist x a* = *dist x b*
shows *a* = *b*
⟨*proof*⟩

lemma *geodesic-segment-union*:
assumes *dist x z* = *dist x y* + *dist y z*
geodesic-segment-between G x y geodesic-segment-between H y z
shows *geodesic-segment-between* (*G* ∪ *H*) *x z*
G ∩ *H* = {*y*}
⟨*proof*⟩

lemma *geodesic-segment-dist-le*:
assumes *geodesic-segment-between G x y a* ∈ *G b* ∈ *G*
shows *dist a b* ≤ *dist x y*
⟨*proof*⟩

lemma *geodesic-segment-param* [*simp*]:
assumes *geodesic-segment-between G x y*
shows *geodesic-segment-param G x 0* = *x*
geodesic-segment-param G x (*dist x y*) = *y*
t ∈ {*0..dist x y*} ⟹ *geodesic-segment-param G x t* ∈ *G*
isometry-on {*0..dist x y*} (*geodesic-segment-param G x*)
(*geodesic-segment-param G x*)‘{*0..dist x y*} = *G*
t ∈ {*0..dist x y*} ⟹ *dist x* (*geodesic-segment-param G x t*) = *t*
s ∈ {*0..dist x y*} ⟹ *t* ∈ {*0..dist x y*} ⟹ *dist* (*geodesic-segment-param G x s*) (*geodesic-segment-param G x t*) = *abs*(*s*−*t*)
z ∈ *G* ⟹ *z* = *geodesic-segment-param G x* (*dist x z*)
⟨*proof*⟩

lemma *geodesic-segment-param-in-segment*:
assumes *G* ≠ {}
shows *geodesic-segment-param G x t* ∈ *G*
⟨*proof*⟩

lemma *geodesic-segment-reverse-param*:
assumes *geodesic-segment-between G x y*
t ∈ {*0..dist x y*}
shows *geodesic-segment-param G y* (*dist x y* − *t*) = *geodesic-segment-param G x t*
⟨*proof*⟩

lemma *dist-along-geodesic-wrt-endpoint*:
assumes *geodesic-segment-between G x y*
u ∈ *G v* ∈ *G*
shows *dist u v* = *abs*(*dist u x* − *dist v x*)

⟨*proof*⟩

One often needs to restrict a geodesic segment to a subsegment. We introduce the tools to express this conveniently.

definition *geodesic-subsegment*::(*'a*::*metric-space*) *set* ⇒ *'a* ⇒ *real* ⇒ *real* ⇒ *'a set*
where *geodesic-subsegment G x s t* = $G \cap \{z.$ *dist x z* $\geq s \wedge$ *dist x z* $\leq t\}$

A subsegment is always contained in the original segment.

lemma *geodesic-subsegment-subset*:
geodesic-subsegment G x s t ⊆ *G*
⟨*proof*⟩

A subsegment is indeed a geodesic segment, and its endpoints and parametrization can be expressed in terms of the original segment.

lemma *geodesic-subsegment*:
assumes *geodesic-segment-between G x y*
$0 \leq s$ $s \leq t$ $t \leq$ *dist x y*
shows *geodesic-subsegment G x s t* = (*geodesic-segment-param G x*)'{*s..t*}
geodesic-segment-between (*geodesic-subsegment G x s t*) (*geodesic-segment-param G x s*) (*geodesic-segment-param G x t*)
$\bigwedge u.$ $s \leq u \Longrightarrow u \leq t \Longrightarrow$ *geodesic-segment-param* (*geodesic-subsegment G x s t*) (*geodesic-segment-param G x s*) $(u - s)$ = *geodesic-segment-param G x u*
⟨*proof*⟩

The parameterizations of a segment and a subsegment sharing an endpoint coincide where defined.

lemma *geodesic-segment-subparam*:
assumes *geodesic-segment-between G x z geodesic-segment-between H x y H* ⊆ *G* *t* ∈ {*0..dist x y*}
shows *geodesic-segment-param G x t* = *geodesic-segment-param H x t*
⟨*proof*⟩

A segment contains a subsegment between any of its points

lemma *geodesic-subsegment-exists*:
assumes *geodesic-segment G x* ∈ *G y* ∈ *G*
shows ∃ *H*. *H* ⊆ *G* ∧ *geodesic-segment-between H x y*
⟨*proof*⟩

A geodesic segment is homeomorphic to an interval.

lemma *geodesic-segment-homeo-interval*:
assumes *geodesic-segment-between G x y*
shows {*0..dist x y*} *homeomorphic G*
⟨*proof*⟩

Just like an interval, a geodesic segment is compact, connected, path connected, bounded, closed, nonempty, and proper.

lemma *geodesic-segment-topology*:
assumes *geodesic-segment G*
shows *compact G connected G path-connected G bounded G closed G* $G \neq \{\}$ *proper G*
⟨*proof*⟩

lemma *geodesic-segment-between-x-x* [*simp*]:
geodesic-segment-between $\{x\}$ *x x*
geodesic-segment $\{x\}$
geodesic-segment-between G x x $\longleftrightarrow G = \{x\}$
⟨*proof*⟩

lemma *geodesic-segment-disconnection*:
assumes *geodesic-segment-between G x y* $z \in G$
shows (*connected* $(G - \{z\})) = (z = x \vee z = y)$
⟨*proof*⟩

lemma *geodesic-segment-unique-endpoints*:
assumes *geodesic-segment-between G x y*
geodesic-segment-between G a b
shows $\{x, y\} = \{a, b\}$
⟨*proof*⟩

lemma *geodesic-segment-subsegment*:
assumes *geodesic-segment G* $H \subseteq G$ *compact H connected H* $H \neq \{\}$
shows *geodesic-segment H*
⟨*proof*⟩

The image under an isometry of a geodesic segment is still obviously a geodesic segment.

lemma *isometry-preserves-geodesic-segment-between*:
assumes *isometry-on X f*
$G \subseteq X$ *geodesic-segment-between G x y*
shows *geodesic-segment-between* (*f‘G*) (*f x*) (*f y*)
⟨*proof*⟩

The sum of distances $d(w, x) + d(w, y)$ can be controlled using the distance from w to a geodesic segment between x and y.

lemma *geodesic-segment-distance*:
assumes *geodesic-segment-between G x y*
shows *dist w x* + *dist w y* $\leq$ *dist x y* + *2* $*$ *infdist w G*
⟨*proof*⟩

If a point y is on a geodesic segment between x and its closest projection p on a set A, then p is also a closest projection of y, and the closest projection set of y is contained in that of x.

lemma *proj-set-geodesic-same-basepoint*:
assumes $p \in$ *proj-set x A geodesic-segment-between G p x* $y \in G$

shows $p \in$ *proj-set y A*
⟨*proof*⟩

lemma *proj-set-subset*:
assumes $p \in$ *proj-set x A geodesic-segment-between G p x* $y \in G$
shows *proj-set y A* $\subseteq$ *proj-set x A*
⟨*proof*⟩

lemma *proj-set-thickening*:
assumes $p \in$ *proj-set x Z*
$0 \leq D$
$D \leq$ *dist p x*
geodesic-segment-between G p x
shows *geodesic-segment-param G p D* $\in$ *proj-set x* ($\bigcup z \in Z$. *cball z D*)
⟨*proof*⟩

lemma *proj-set-thickening'*:
assumes $p \in$ *proj-set x Z*
$0 \leq D$
$D \leq E$
$E \leq$ *dist p x*
geodesic-segment-between G p x
shows *geodesic-segment-param G p D* $\in$ *proj-set* (*geodesic-segment-param G p E*) ($\bigcup z \in Z$. *cball z D*)
⟨*proof*⟩

It is often convenient to use *one* geodesic between x and y, even if it is not unique. We introduce a notation for such a choice of a geodesic, denoted `{x--S--y}` for such a geodesic that moreover remains in the set S. We also enforce the condition `{x--S--y} = {y--S--x}`. When there is no such geodesic, we simply take `{x--S--y} = {x, y}` for definiteness. It would be even better to enforce that, if a is on `{x--S--y}`, then `{x--S--y}` is the union of `{x--S--a}` and `{a--S--y}`, but I do not know if such a choice is always possible – such a choice of geodesics is called a geodesic bicombing. We also write `{x--y}` for `{x--UNIV--y}`.

definition *some-geodesic-segment-between*::'*a*::*metric-space* ⇒ '*a set* ⇒ '*a* ⇒ '*a set* (‹(*1*{-−−-−−-})›)
where *some-geodesic-segment-between* = (*SOME f*. ∀ *x y S*. *f x S y* = *f y S x*
∧ (*if* (∃ *G*. *geodesic-segment-between G x y* ∧ *G* ⊆ *S*) *then* (*geodesic-segment-between* (*f x S y*) *x y* ∧ (*f x S y* ⊆ *S*))
else f x S y = {*x*, *y*}))

abbreviation *some-geodesic-segment-between-UNIV*::'*a*::*metric-space* ⇒ '*a* ⇒ '*a set* (‹(*1*{-−−-})›)
where *some-geodesic-segment-between-UNIV x y* ≡ {*x*−−*UNIV*−−*y*}

We prove that there is such a choice of geodesics, compatible with direction reversal. What we do is choose arbitrarily a geodesic between x and y if it exists, and then use the geodesic between $\min(x, y)$ and $\max(x, y)$, for any total order on the space, to ensure that we get the same result from x to y or from y to x.

lemma *some-geodesic-segment-between-exists*:
$\exists f.\ \forall\ x\ y\ S.\ f\ x\ S\ y = f\ y\ S\ x$
$\wedge$ (*if* ($\exists$ *G*. *geodesic-segment-between* $G\ x\ y \wedge G \subseteq S$) *then* (*geodesic-segment-between* $(f\ x\ S\ y)\ x\ y \wedge (f\ x\ S\ y \subseteq S)$)
else $f\ x\ S\ y = \{x, y\}$)
⟨*proof*⟩

lemma *some-geodesic-commute*:
$\{x{-}{-}S{-}{-}y\} = \{y{-}{-}S{-}{-}x\}$
⟨*proof*⟩

lemma *some-geodesic-segment-description*:
($\exists$ *G*. *geodesic-segment-between* $G\ x\ y \wedge G \subseteq S$) $\Longrightarrow$ *geodesic-segment-between* $\{x{-}{-}S{-}{-}y\}\ x\ y$
($\neg$($\exists$ *G*. *geodesic-segment-between* $G\ x\ y \wedge G \subseteq S$)) $\Longrightarrow \{x{-}{-}S{-}{-}y\} = \{x, y\}$
⟨*proof*⟩

Basic topological properties of our chosen set of geodesics.

lemma *some-geodesic-compact* [*simp*]:
compact $\{x{-}{-}S{-}{-}y\}$
⟨*proof*⟩

lemma *some-geodesic-closed* [*simp*]:
closed $\{x{-}{-}S{-}{-}y\}$
⟨*proof*⟩

lemma *some-geodesic-bounded* [*simp*]:
bounded $\{x{-}{-}S{-}{-}y\}$
⟨*proof*⟩

lemma *some-geodesic-endpoints* [*simp*]:
$x \in \{x{-}{-}S{-}{-}y\}$ $y \in \{x{-}{-}S{-}{-}y\}$ $\{x{-}{-}S{-}{-}y\} \neq \{\}$
⟨*proof*⟩

lemma *some-geodesic-subsegment*:
assumes $H \subseteq \{x{-}{-}S{-}{-}y\}$ *compact H connected H* $H \neq \{\}$
shows *geodesic-segment H*
⟨*proof*⟩

lemma *some-geodesic-in-subset*:
assumes $x \in S$ $y \in S$
shows $\{x{-}{-}S{-}{-}y\} \subseteq S$
⟨*proof*⟩

lemma *some-geodesic-same-endpoints* [*simp*]:
$\{x{-}{-}S{-}{-}x\} = \{x\}$

⟨*proof*⟩

5.2 Geodesic subsets

A subset is *geodesic* if any two of its points can be joined by a geodesic segment. We prove basic properties of such a subset in this paragraph – notably connectedness. A basic example is given by convex subsets of vector spaces, as closed segments are geodesic.

definition *geodesic-subset*::(*'a*::*metric-space*) *set* ⇒ *bool*
where *geodesic-subset S* = (∀ *x*∈*S*. ∀ *y*∈*S*. ∃ *G*. *geodesic-segment-between G x y* ∧ *G* ⊆ *S*)

lemma *geodesic-subsetD*:
assumes *geodesic-subset S x* ∈ *S y* ∈ *S*
shows *geodesic-segment-between* {*x*−−*S*−−*y*} *x y*
⟨*proof*⟩

lemma *geodesic-subsetI*:
assumes ⋀*x y*. *x* ∈ *S* ⟹ *y* ∈ *S* ⟹ ∃ *G*. *geodesic-segment-between G x y* ∧ *G* ⊆ *S*
shows *geodesic-subset S*
⟨*proof*⟩

lemma *geodesic-subset-empty*:
geodesic-subset {}
⟨*proof*⟩

lemma *geodesic-subset-singleton*:
geodesic-subset {*x*}
⟨*proof*⟩

lemma *geodesic-subset-path-connected*:
assumes *geodesic-subset S*
shows *path-connected S*
⟨*proof*⟩

To show that a segment in a normed vector space is geodesic, we will need to use its length parametrization, which is given in the next lemma.

lemma *closed-segment-as-isometric-image*:
((λ*t*. *x* + (*t*/*dist x y*) $*_R$ (*y* − *x*))‘{*0*..*dist x y*}) = *closed-segment x y*
⟨*proof*⟩

proposition *closed-segment-is-geodesic*:
fixes *x y*::*'a*::*real-normed-vector*
shows *isometry-on* {*0*..*dist x y*} (λ*t*. *x* + (*t*/*dist x y*) $*_R$ (*y* − *x*))
geodesic-segment-between (*closed-segment x y*) *x y*
geodesic-segment (*closed-segment x y*)
⟨*proof*⟩

We deduce that a convex set is geodesic.

proposition *convex-is-geodesic*:
  assumes *convex* (*S*::*'a*::*real-normed-vector set*)
  shows *geodesic-subset S*
⟨*proof*⟩

5.3 Geodesic spaces

In this subsection, we define geodesic spaces (metric spaces in which there is a geodesic segment joining any pair of points). We specialize the previous statements on geodesic segments to these situations.

class *geodesic-space* = *metric-space* +
  assumes *geodesic*: *geodesic-subset* (*UNIV*::(*'a*::*metric-space*) *set*)

The simplest example of a geodesic space is a real normed vector space. Significant examples also include graphs (with the graph distance), Riemannian manifolds, and $CAT(\kappa)$ spaces.

instance *real-normed-vector* ⊆ *geodesic-space*
⟨*proof*⟩

lemma (**in** *geodesic-space*) *some-geodesic-is-geodesic-segment* [*simp*]:
  geodesic-segment-between {*x*−−*y*} *x* (*y*::*'a*)
  geodesic-segment {*x*−−*y*}
⟨*proof*⟩

lemma (**in** *geodesic-space*) *some-geodesic-connected* [*simp*]:
  connected {*x*−−*y*} *path-connected* {*x*−−*y*}
⟨*proof*⟩

In geodesic spaces, we restate as simp rules all properties of the geodesic segment parametrizations.

lemma (**in** *geodesic-space*) *geodesic-segment-param-in-geodesic-spaces* [*simp*]:
  geodesic-segment-param {*x*−−*y*} *x 0* = *x*
  geodesic-segment-param {*x*−−*y*} *x* (*dist x y*) = *y*
  t ∈ {*0..dist x y*} ⟹ *geodesic-segment-param* {*x*−−*y*} *x t* ∈ {*x*−−*y*}
  isometry-on {*0..dist x y*} (*geodesic-segment-param* {*x*−−*y*} *x*)
  (*geodesic-segment-param* {*x*−−*y*} *x*)'{*0..dist x y*} = {*x*−−*y*}
  t ∈ {*0..dist x y*} ⟹ *dist x* (*geodesic-segment-param* {*x*−−*y*} *x t*) = *t*
  s ∈ {*0..dist x y*} ⟹ *t* ∈ {*0..dist x y*} ⟹ *dist* (*geodesic-segment-param* {*x*−−*y*} *x s*) (*geodesic-segment-param* {*x*−−*y*} *x t*) = *abs*(*s*−*t*)
  z ∈ {*x*−−*y*} ⟹ *z* = *geodesic-segment-param* {*x*−−*y*} *x* (*dist x z*)
⟨*proof*⟩

5.4 Uniquely geodesic spaces

In this subsection, we define uniquely geodesic spaces, i.e., geodesic spaces in which, additionally, there is a unique geodesic between any pair of points.

class *uniquely-geodesic-space* = *geodesic-space* +
assumes *uniquely-geodesic*: $\bigwedge x\ y\ G\ H.$ *geodesic-segment-between* $G\ x\ y \Longrightarrow$ *geodesic-segment-between* $H\ x\ y \Longrightarrow G = H$

To prove that a geodesic space is uniquely geodesic, it suffices to show that there is no loop, i.e., if two geodesic segments intersect only at their endpoints, then they coincide.

Indeed, assume this holds, and consider two geodesics with the same endpoints. If they differ at some time t, then consider the last time a before t where they coincide, and the first time b after t where they coincide. Then the restrictions of the two geodesics to $[a, b]$ give a loop, and a contradiction.

lemma (**in** *geodesic-space*) *uniquely-geodesic-spaceI*:
assumes $\bigwedge G\ H\ x\ (y{::}'a).$ *geodesic-segment-between* $G\ x\ y \Longrightarrow$ *geodesic-segment-between* $H\ x\ y \Longrightarrow G \cap H = \{x, y\} \Longrightarrow x = y$
geodesic-segment-between $G\ x\ y$ *geodesic-segment-between* $H\ x\ (y{::}'a)$
shows $G = H$
⟨*proof*⟩

context *uniquely-geodesic-space*
begin

lemma *geodesic-segment-unique*:
geodesic-segment-between $G\ x\ y = (G = \{x{-}{-}(y{::}'a)\})$
⟨*proof*⟩

lemma *geodesic-segment-dist'*:
assumes *dist* $x\ z$ = *dist* $x\ y$ + *dist* $y\ z$
shows $y \in \{x{-}{-}z\}$ $\{x{-}{-}z\} = \{x{-}{-}y\} \cup \{y{-}{-}z\}$
⟨*proof*⟩

lemma *geodesic-segment-expression*:
$\{x{-}{-}z\} = \{y.$ *dist* $x\ z$ = *dist* $x\ y$ + *dist* $y\ z\}$
⟨*proof*⟩

lemma *geodesic-segment-split*:
assumes $(y{::}'a) \in \{x{-}{-}z\}$
shows $\{x{-}{-}z\} = \{x{-}{-}y\} \cup \{y{-}{-}z\}$
$\{x{-}{-}y\} \cap \{y{-}{-}z\} = \{y\}$
⟨*proof*⟩

lemma *geodesic-segment-subparam'*:
assumes $y \in \{x{-}{-}z\}$ $t \in \{0..dist\ x\ y\}$
shows *geodesic-segment-param* $\{x{-}{-}z\}\ x\ t$ = *geodesic-segment-param* $\{x{-}{-}y\}\ x\ t$
⟨*proof*⟩

end

5.5 A complete metric space with middles is geodesic.

A complete space in which every pair of points has a middle (i.e., a point m which is half distance of x and y) is geodesic: to construct a geodesic between x_0 and y_0, first choose a middle m, then middles of the pairs (x_0, m) and (m, y_0), and so on. This will define the geodesic on dyadic points (and this is indeed an isometry on these dyadic points. Then, extend it by uniform continuity to the whole segment $[0, dist x0y0]$.

The formal proof will be done in a locale where x_0 and y_0 are fixed, for notational simplicity. We define inductively the sequence of middles, in a function `geod` of two natural variables: *geodnm* corresponds to the image of the dyadic point $m/2^n$. It is defined inductively, by $geod(n+1)(2m) = geodnm$, and $geod(n+1)(2m+1)$ is a middle of *geodnm* and $geodn(m+1)$. This is not a completely classical inductive definition, so one has to use `function` to define it. Then, one checks inductively that it has all the properties we want, and use it to define the geodesic segment on dyadic points. We will not use a canonical representative for a dyadic point, but any representative (i.e., numerator and denominator will not have to be coprime) – this will not create problems as *geod* does not depend on the choice of the representative, by construction.

locale *complete-space-with-middle* =
  fixes *x0 y0*::*'a*::*complete-space*
  assumes *middles*: $\bigwedge x\ y::'a.\ \exists z.\ dist\ x\ z = (dist\ x\ y)/2 \wedge dist\ z\ y = (dist\ x\ y)/2$
begin

definition *middle*::*'a ⇒ 'a ⇒ 'a*
  where *middle x y = (SOME z. dist x z = (dist x y)/2 ∧ dist z y = (dist x y)/2)*

lemma *middle*:
  dist x (middle x y) = (dist x y)/2
  dist (middle x y) y = (dist x y)/2
⟨*proof*⟩

function *geod*::*nat ⇒ nat ⇒ 'a* **where**
 geod 0 0 = x0
|*geod 0 (Suc m) = y0*
|*geod (Suc n) (2 * m) = geod n m*
|*geod (Suc n) (Suc (2*m)) = middle (geod n m) (geod n (Suc m))*
⟨*proof*⟩
termination ⟨*proof*⟩

By induction, the distance between successive points is $D/2^n$.

lemma *geod-distance-successor*:
  ∀ a < 2^n. dist (geod n a) (geod n (Suc a)) = dist x0 y0 / 2^n
⟨*proof*⟩

lemma *geod-mult*:

geod n a = geod (n + k) (a ∗ 2^k)
⟨*proof*⟩

lemma *geod-0*:
geod n 0 = x0
⟨*proof*⟩

lemma *geod-end*:
geod n (2^n) = y0
⟨*proof*⟩

By the triangular inequality, the distance between points separated by $(b - a)/2^n$ is at most $D * (b - a)/2^n$.

lemma *geod-upper*:
assumes *a ≤ b b ≤ 2^n*
shows *dist (geod n a) (geod n b) ≤ (b−a) ∗ dist x0 y0 / 2^n*
⟨*proof*⟩

In fact, the distance is exactly $D * (b - a)/2^n$, otherwise the extremities of the interval would be closer than D, a contradiction.

lemma *geod-dist*:
assumes *a ≤ b b ≤ 2^n*
shows *dist (geod n a) (geod n b) = (b−a) ∗ dist x0 y0 / 2^n*
⟨*proof*⟩

We deduce the same statement but for points that are not on the same level, by putting them on a common multiple level.

lemma *geod-dist2*:
assumes *a ≤ 2^n b ≤ 2^p a/2^n ≤ b / 2^p*
shows *dist (geod n a) (geod p b) = (b/2^p − a/2^n) ∗ dist x0 y0*
⟨*proof*⟩

Same thing but without a priori ordering of the points.

lemma *geod-dist3*:
assumes *a ≤ 2^n b ≤ 2^p*
shows *dist (geod n a) (geod p b) = abs(b/2^p − a/2^n) ∗ dist x0 y0*
⟨*proof*⟩

Finally, we define a geodesic by extending what we have already defined on dyadic points, thanks to the result of isometric extension of isometries taking their values in complete spaces.

lemma *geod*:
shows *∃ g. isometry-on {0..dist x0 y0} g ∧ g 0 = x0 ∧ g (dist x0 y0) = y0*
⟨*proof*⟩

end

We can now complete the proof that a complete space with middles is in fact geodesic: all the work has been done in the locale `complete_space_with_middle`, in Lemma `geod`.

theorem *complete-with-middles-imp-geodesic*:
assumes $\bigwedge x\ y$::(*'a*::*complete-space*). $\exists\, m.\ dist\ x\ m = dist\ x\ y\ /2 \wedge dist\ m\ y = dist\ x\ y\ /2$
shows *OFCLASS*(*'a*, *geodesic-space-class*)
⟨*proof*⟩

6 Quasi-isometries

A (λ, C) quasi-isometry is a function which behaves like an isometry, up to an additive error C and a multiplicative error λ. It can be very different from an isometry on small scales (for instance, the function integer part is a quasi-isometry between $\mathbb{R}$ and $\mathbb{Z}$), but on large scales it captures many important features of isometries.

When the space is unbounded, one checks easily that $C \geq 0$ and $\lambda \geq 1$. As this is the only case of interest (any two bounded sets are quasi-isometric), we incorporate this requirement in the definition.

definition *quasi-isometry-on*::*real* ⇒ *real* ⇒ (*'a*::*metric-space*) *set* ⇒ (*'a* ⇒ (*'b*::*metric-space*)) ⇒ *bool*
(‹- - −*quasi'-isometry'-on*› [*1000*, *999*])
where $lambda\ C{-}quasi\text{-}isometry\text{-}on\ X\ f = ((lambda \geq 1) \wedge (C \geq 0) \wedge (\forall x \in X.\ \forall y \in X.\ (dist\ (f\ x)\ (f\ y) \leq lambda * dist\ x\ y + C \wedge dist\ (f\ x)\ (f\ y) \geq (1/lambda) * dist\ x\ y - C)))$

abbreviation *quasi-isometry* :: *real* ⇒ *real* ⇒ (*'a*::*metric-space* ⇒ *'b*::*metric-space*) ⇒ *bool*
(‹- - −*quasi'-isometry*› [*1000*, *999*])
where *quasi-isometry lambda C f* ≡ *lambda C*−*quasi-isometry-on UNIV f*

6.1 Basic properties of quasi-isometries

lemma *quasi-isometry-onD*:
assumes *lambda C*−*quasi-isometry-on X f*
shows $\bigwedge x\ y.\ x \in X \Longrightarrow y \in X \Longrightarrow dist\ (f\ x)\ (f\ y) \leq lambda * dist\ x\ y + C$
$\bigwedge x\ y.\ x \in X \Longrightarrow y \in X \Longrightarrow dist\ (f\ x)\ (f\ y) \geq (1/lambda) * dist\ x\ y - C$
$lambda \geq 1\ C \geq 0$
⟨*proof*⟩

lemma *quasi-isometry-onI* [*intro*]:
assumes $\bigwedge x\ y.\ x \in X \Longrightarrow y \in X \Longrightarrow dist\ (f\ x)\ (f\ y) \leq lambda * dist\ x\ y + C$
$\bigwedge x\ y.\ x \in X \Longrightarrow y \in X \Longrightarrow dist\ (f\ x)\ (f\ y) \geq (1/lambda) * dist\ x\ y - C$
$lambda \geq 1\ C \geq 0$
shows *lambda C*−*quasi-isometry-on X f*
⟨*proof*⟩

lemma *isometry-quasi-isometry-on*:
  assumes *isometry-on X f*
  shows *1 0−quasi-isometry-on X f*
⟨*proof*⟩

lemma *quasi-isometry-on-change-params*:
  assumes *lambda C−quasi-isometry-on X f mu* $\geq$ *lambda D* $\geq$ *C*
  shows *mu D−quasi-isometry-on X f*
⟨*proof*⟩

lemma *quasi-isometry-on-subset*:
  assumes *lambda C−quasi-isometry-on X f*
    $Y \subseteq X$
  shows *lambda C−quasi-isometry-on Y f*
⟨*proof*⟩

lemma *quasi-isometry-on-perturb*:
  assumes *lambda C−quasi-isometry-on X f*
    $D \geq 0$
    $\bigwedge x.\ x \in X \Longrightarrow dist\ (f\ x)\ (g\ x) \leq D$
  shows *lambda (C + 2 ∗ D)−quasi-isometry-on X g*
⟨*proof*⟩

lemma *quasi-isometry-on-compose*:
  assumes *lambda C−quasi-isometry-on X f*
    mu D−quasi-isometry-on Y g
    f'X $\subseteq$ *Y*
  shows *(lambda ∗ mu) (C ∗ mu + D)−quasi-isometry-on X (g o f)*
⟨*proof*⟩

lemma *quasi-isometry-on-bounded*:
  assumes *lambda C−quasi-isometry-on X f*
    bounded X
  shows *bounded (f'X)*
⟨*proof*⟩

lemma *quasi-isometry-on-empty*:
  assumes $C \geq 0$ *lambda* ≥ 1
  shows *lambda C−quasi-isometry-on {} f*
⟨*proof*⟩

Quasi-isometries change the distance to a set by at most $\lambda \cdot +C$, this follows readily from the fact that this inequality holds pointwise.

lemma *quasi-isometry-on-infdist*:
  assumes *lambda C−quasi-isometry-on X f*
    $w \in X$
    $S \subseteq X$
  shows *infdist (f w) (f'S)* $\leq$ *lambda ∗ infdist w S + C*

infdist (*f w*) (*f'S*) $\geq$ (*1*/*lambda*) $*$ *infdist w S* $-$ *C*
⟨*proof*⟩

6.2 Quasi-isometric isomorphisms

The notion of isomorphism for quasi-isometries is not that it should be a bijection, as it is a coarse notion, but that it is a bijection up to a bounded displacement. For instance, the inclusion of $\mathbb{Z}$ in $\mathbb{R}$ is a quasi-isometric isomorphism between these spaces, whose (quasi)-inverse (which is non-unique) is given by the function integer part. This is formalized in the next definition.

definition *quasi-isometry-between*::*real* ⇒ *real* ⇒ (*'a*::*metric-space*) *set* ⇒ (*'b*::*metric-space*) *set* ⇒ (*'a* ⇒ *'b*) ⇒ *bool*
(‹- - −*quasi'-isometry'-between*› [*1000*, *999*])
where *lambda C*−*quasi-isometry-between X Y f* = ((*lambda C*−*quasi-isometry-on X f*) ∧ (*f'X* ⊆ *Y*) ∧ (∀ *y*∈*Y*. ∃ *x*∈*X*. *dist* (*f x*) *y* ≤ *C*))

definition *quasi-isometric*::(*'a*::*metric-space*) *set* ⇒ (*'b*::*metric-space*) *set* ⇒ *bool*
where *quasi-isometric X Y* = (∃ *lambda C f*. *lambda C*−*quasi-isometry-between X Y f*)

lemma *quasi-isometry-betweenD*:
assumes *lambda C*−*quasi-isometry-between X Y f*
shows *lambda C*−*quasi-isometry-on X f*
f'X ⊆ *Y*
⋀*y*. *y* ∈ *Y* ⟹ ∃ *x*∈*X*. *dist* (*f x*) *y* ≤ *C*
⋀*x y*. *x* ∈ *X* ⟹ *y* ∈ *X* ⟹ *dist* (*f x*) (*f y*) ≤ *lambda* ∗ *dist x y* + *C*
⋀*x y*. *x* ∈ *X* ⟹ *y* ∈ *X* ⟹ *dist* (*f x*) (*f y*) ≥ (*1*/*lambda*) ∗ *dist x y* − *C*
lambda ≥ *1 C* ≥ *0*
⟨*proof*⟩

lemma *quasi-isometry-betweenI*:
assumes *lambda C*−*quasi-isometry-on X f*
f'X ⊆ *Y*
⋀*y*. *y* ∈ *Y* ⟹ ∃ *x*∈*X*. *dist* (*f x*) *y* ≤ *C*
shows *lambda C*−*quasi-isometry-between X Y f*
⟨*proof*⟩

lemma *quasi-isometry-on-between*:
assumes *lambda C*−*quasi-isometry-on X f*
shows *lambda C*−*quasi-isometry-between X* (*f'X*) *f*
⟨*proof*⟩

lemma *quasi-isometry-between-change-params*:
assumes *lambda C*−*quasi-isometry-between X Y f mu* ≥ *lambda D* ≥ *C*
shows *mu D*−*quasi-isometry-between X Y f*
⟨*proof*⟩

lemma *quasi-isometry-subset*:
assumes $X \subseteq Y$ $\bigwedge y.\ y \in Y \Longrightarrow \exists x{\in}X.\ dist\ x\ y \leq C$ $C \geq 0$
shows $1\ C{-}quasi\text{-}isometry\text{-}between\ X\ Y\ (\lambda x.\ x)$
⟨*proof*⟩

lemma *isometry-quasi-isometry-between*:
assumes *isometry f*
shows $1\ 0{-}quasi\text{-}isometry\text{-}between\ UNIV\ UNIV\ f$
⟨*proof*⟩

proposition *quasi-isometry-inverse*:
assumes $lambda\ C{-}quasi\text{-}isometry\text{-}between\ X\ Y\ f$
shows $\exists g.\ lambda\ (3 * C * lambda){-}quasi\text{-}isometry\text{-}between\ Y\ X\ g$
$\wedge (\forall x{\in}X.\ dist\ x\ (g\ (f\ x)) \leq 3 * C * lambda)$
$\wedge (\forall y{\in}Y.\ dist\ y\ (f\ (g\ y)) \leq 3 * C * lambda)$
⟨*proof*⟩

proposition *quasi-isometry-compose*:
assumes $lambda\ C{-}quasi\text{-}isometry\text{-}between\ X\ Y\ f$
$mu\ D{-}quasi\text{-}isometry\text{-}between\ Y\ Z\ g$
shows $(lambda * mu)\ (C * mu + 2 * D){-}quasi\text{-}isometry\text{-}between\ X\ Z\ (g\ o\ f)$
⟨*proof*⟩

theorem *quasi-isometric-equiv-rel*:
quasi-isometric X X
$quasi\text{-}isometric\ X\ Y \Longrightarrow quasi\text{-}isometric\ Y\ Z \Longrightarrow quasi\text{-}isometric\ X\ Z$
$quasi\text{-}isometric\ X\ Y \Longrightarrow quasi\text{-}isometric\ Y\ X$
⟨*proof*⟩

Many interesting properties in geometric group theory are invariant under quasi-isometry. We prove the most basic ones here.

lemma *quasi-isometric-empty*:
assumes $X = \{\}$ *quasi-isometric X Y*
shows $Y = \{\}$
⟨*proof*⟩

lemma *quasi-isometric-bounded*:
assumes *bounded X* *quasi-isometric X Y*
shows *bounded Y*
⟨*proof*⟩

lemma *quasi-isometric-bounded-iff*:
assumes *bounded X* $X \neq \{\}$ *bounded Y* $Y \neq \{\}$
shows *quasi-isometric X Y*
⟨*proof*⟩

6.3 Quasi-isometries of Euclidean spaces.

A less trivial fact is that the dimension of euclidean spaces is invariant under quasi-isometries. It is proved below using growth argument, as quasi-isometries preserve the growth rate.

The growth of the space is asymptotic behavior of the number of well-separated points that fit in a ball of radius R, when R tends to infinity. Up to a suitable equivalence, it is clearly a quasi-isometry invariance. We show below that, in a Euclidean space of dimension d, the growth is like R^d: the upper bound is obtained by using the fact that we have disjoint balls inside a big ball, hence volume controls conclude the argument, while the lower bound is obtained by considering integer points.

First, we show that the growth rate of a Euclidean space of dimension d is bounded from above by R^d, using the control on measure of disjoint balls and a volume argument.

proposition *growth-rate-euclidean-above*:
  fixes $D::real$
  assumes $D > (0::real)$
    and H: $F \subseteq cball\ (0::'a::euclidean\text{-}space)\ R\ R \geq 0$
      $\bigwedge x\ y.\ x \in F \Longrightarrow y \in F \Longrightarrow x \neq y \Longrightarrow dist\ x\ y \geq D$
  shows $finite\ F \wedge card\ F \leq 1 + ((6/D)\widehat{\ }(DIM('a))) * R\widehat{\ }(DIM('a))$
$\langle proof \rangle$

Then, we show that the growth rate of a Euclidean space of dimension d is bounded from below by R^d, using integer points.

proposition *growth-rate-euclidean-below*:
  fixes $D::real$
  assumes $R \geq 0$
  shows $\exists F.\ (F \subseteq cball\ (0::'a::euclidean\text{-}space)\ R$
      $\wedge (\forall x{\in}F.\ \forall y{\in}F.\ x = y \vee dist\ x\ y \geq D) \wedge finite\ F \wedge card\ F \geq (1/((max$
$D\ 1) * DIM('a)))\widehat{\ }(DIM('a)) * R\widehat{\ }(DIM('a)))$
$\langle proof \rangle$

As the growth is invariant under quasi-isometries, we deduce that it is impossible to map quasi-isometrically a Euclidean space in a space of strictly smaller dimension.

proposition *quasi-isometry-on-euclidean*:
  fixes $f::'a::euclidean\text{-}space \Rightarrow 'b::euclidean\text{-}space$
  assumes $lambda\ C{-}quasi\text{-}isometry\text{-}on\ UNIV\ f$
  shows $DIM('a) \leq DIM('b)$
$\langle proof \rangle$

As a particular case, we deduce that two quasi-isometric Euclidean spaces have the same dimension.

theorem *quasi-isometric-euclidean*:

assumes *quasi-isometric* (*UNIV*::*'a*::*euclidean-space set*) (*UNIV*::*'b*::*euclidean-space set*)
shows *DIM*(*'a*) = *DIM*(*'b*)
⟨*proof*⟩

A different (and important) way to prove the above statement would be to use asymptotic cones. Here, it can be done in an elementary way: start with a quasi-isometric map f, and consider a limit (defined with a ultrafilter) of $x \mapsto f(nx)/n$. This is a map which contracts and expands the distances by at most λ. In particular, it is a homeomorphism on its image. No such map exists if the dimension of the target is smaller than the dimension of the source (invariance of domain theorem, already available in the library).

The above argument using growth is more elementary to write, though.

6.4 Quasi-geodesics

A quasi-geodesic is a quasi-isometric embedding of a real segment into a metric space. As the embedding need not be continuous, a quasi-geodesic does not have to be compact, nor connected, which can be a problem. However, in a geodesic space, it is always possible to deform a quasi-geodesic into a continuous one (at the price of worsening the quasi-isometry constants). This is the content of the proposition `quasi_geodesic_made_lipschitz` below, which is a variation around Lemma III.H.1.11 in [BH99]. The strategy of the proof is simple: assume that the quasi-geodesic c is defined on $[a, b]$. Then, on the points a, $a + C/\lambda$, $\cdots$, $a + N \cdot C/\lambda$, b, take d equal to c, where N is chosen so that the distance between the last point and b is in $[C/\lambda, 2C/\lambda)$. In the intervals, take d to be geodesic.

proposition (**in** *geodesic-space*) *quasi-geodesic-made-lipschitz*:
fixes *c*::*real* ⇒ *'a*
assumes *lambda C−quasi-isometry-on* {*a*..*b*} *c dist* (*c a*) (*c b*) ≥ *2* ∗ *C*
shows ∃ *d*. *continuous-on* {*a*..*b*} *d* ∧ *d a* = *c a* ∧ *d b* = *c b*
∧ (∀ *x*∈{*a*..*b*}. *dist* (*c x*) (*d x*) ≤ *4* ∗ *C*)
∧ *lambda* (*4* ∗ *C*)−*quasi-isometry-on* {*a*..*b*} *d*
∧ (*2* ∗ *lambda*)−*lipschitz-on* {*a*..*b*} *d*
∧ *hausdorff-distance* (*c*‘{*a*..*b*}) (*d*‘{*a*..*b*}) ≤ *2* ∗ *C*
⟨*proof*⟩

end

7 The metric completion of a metric space

theory *Metric-Completion*
imports *Isometries*
begin

Any metric space can be completed, by adding the missing limits of Cauchy sequences. Formally, there exists an isometric embedding of the space in a complete space, with dense image. In this paragraph, we construct this metric completion. This is exactly the same construction as the way in which real numbers are constructed from rational numbers.

7.1 Definition of the metric completion

quotient-type (**overloaded**) *'a metric-completion* =
nat ⇒ (*'a::metric-space*) / *partial*: $\lambda u\ v.\ (Cauchy\ u) \wedge (Cauchy\ v) \wedge (\lambda n.\ dist\ (u\ n)\ (v\ n)) \longrightarrow 0$
⟨*proof*⟩

We have to show that the metric completion is indeed a metric space, that the original space embeds isometrically into it, and that it is complete. Before we prove these statements, we start with two simple lemmas that will be needed later on.

lemma *convergent-Cauchy-dist*:
fixes *u v::nat* ⇒ (*'a::metric-space*)
assumes *Cauchy u Cauchy v*
shows *convergent* ($\lambda n.\ dist\ (u\ n)\ (v\ n)$)
⟨*proof*⟩

lemma *convergent-add-null*:
fixes *u v::nat* ⇒ (*'a::real-normed-vector*)
assumes *convergent u*
$(\lambda n.\ v\ n - u\ n) \longrightarrow 0$
shows *convergent v lim v = lim u*
⟨*proof*⟩

Let us now prove that the metric completion is a metric space: the distance between two Cauchy sequences is the limit of the distances of points in the sequence. The convergence follows from Lemma `convergent_Cauchy_dist` above.

instantiation *metric-completion* :: (*metric-space*) *metric-space*
begin

lift-definition *dist-metric-completion*::(*'a::metric-space*) *metric-completion* ⇒ *'a metric-completion* ⇒ *real*
is $\lambda x\ y.\ lim\ (\lambda n.\ dist\ (x\ n)\ (y\ n))$
⟨*proof*⟩

lemma *dist-metric-completion-limit*:
fixes *x y::'a metric-completion*
shows ($\lambda n.\ dist$ (*rep-metric-completion x n*) (*rep-metric-completion y n*)) ⟶ *dist x y*
⟨*proof*⟩

lemma *dist-metric-completion-limit'*:
  fixes *x y*::*nat* ⇒ *'a*
  assumes *Cauchy x Cauchy y*
  shows ($\lambda n.$ *dist* (*x n*) (*y n*)) ⟶ *dist* (*abs-metric-completion x*) (*abs-metric-completion y*)
⟨*proof*⟩

To define a metric space in the current library of Isabelle/HOL, one should also introduce a uniformity structure and a topology, as follows (they are prescribed by the distance):

definition *uniformity-metric-completion*::((*'a metric-completion*) × (*'a metric-completion*)) *filter*
  where *uniformity-metric-completion* = (*INF e*∈{*0* <..}. *principal* {(*x*, *y*). *dist x y* < *e*})

definition *open-metric-completion* :: *'a metric-completion set* ⇒ *bool*
  where *open-metric-completion U* = (∀ *x*∈*U*. *eventually* (λ(*x'*, *y*). *x'* = *x* ⟶ *y* ∈ *U*) *uniformity*)

instance ⟨*proof*⟩
end

Let us now show that the distance thus defined on the metric completion is indeed complete. This is essentially by design.

instance *metric-completion* :: (*metric-space*) *complete-space*
⟨*proof*⟩

7.2 Isometric embedding of a space in its metric completion

The canonical embedding of a space into its metric completion is obtained by taking the Cauchy sequence which is constant, equal to the given point. This is indeed an isometric embedding with dense image, as we prove in the lemmas below.

definition *to-metric-completion*::(*'a*::*metric-space*) ⇒ *'a metric-completion*
  where *to-metric-completion x* = *abs-metric-completion* (λ*n*. *x*)

lemma *to-metric-completion-isometry*:
  isometry-on UNIV to-metric-completion
⟨*proof*⟩

lemma *to-metric-completion-dense*:
  assumes *open U U* ≠ {}
  shows ∃ *x*. *to-metric-completion x* ∈ *U*
⟨*proof*⟩

lemma *to-metric-completion-dense'*:
  closure (*range to-metric-completion*) = *UNIV*

⟨*proof*⟩

The main feature of the completion is that a uniformly continuous function on the original space can be extended to a uniformly continuous function on the completion, i.e., it can be written as the composition of a new function and of the inclusion `to_metric_completion`.

lemma *lift-to-metric-completion*:
  fixes $f::('a::metric\text{-}space) \Rightarrow ('b::complete\text{-}space)$
  assumes *uniformly-continuous-on UNIV f*
  shows $\exists\, g.\ (uniformly\text{-}continuous\text{-}on\ UNIV\ g)$
    $\wedge\ (f = g\ o\ to\text{-}metric\text{-}completion)$
    $\wedge\ (\forall\, x \in range\ to\text{-}metric\text{-}completion.\ g\ x = f\ (inv\ to\text{-}metric\text{-}completion$
$x))$
⟨*proof*⟩

When the function is an isometry, the lifted function is also an isometry (and its range is the closure of the range of the original function). This shows that the metric completion is unique, up to isometry:

lemma *lift-to-metric-completion-isometry*:
  fixes $f::('a::metric\text{-}space) \Rightarrow ('b::complete\text{-}space)$
  assumes *isometry-on UNIV f*
  shows $\exists\, g.\ isometry\text{-}on\ UNIV\ g$
    $\wedge\ range\ g = closure(range\ f)$
    $\wedge\ f = g\ o\ to\text{-}metric\text{-}completion$
    $\wedge\ (\forall\, x \in range\ to\text{-}metric\text{-}completion.\ g\ x = f\ (inv\ to\text{-}metric\text{-}completion\ x))$
⟨*proof*⟩

7.3 The metric completion of a second countable space is second countable

We want to show that the metric completion of a second countable space is still second countable. This is most easily expressed using the fact that a metric space is second countable if and only if there exists a dense countable subset. We prove the equivalence in the next lemma, and use it then to prove that the metric completion is still second countable.

lemma *second-countable-iff-dense-countable-subset*:
  $(\exists\, B::'a::metric\text{-}space\ set\ set.\ countable\ B \wedge topological\text{-}basis\ B)$
    $\longleftrightarrow (\exists\, A::'a\ set.\ countable\ A \wedge closure\ A = UNIV)$
⟨*proof*⟩

lemma *second-countable-metric-dense-subset*:
  $\exists\, A::'a::\{metric\text{-}space, second\text{-}countable\text{-}topology\}\ set.\ countable\ A \wedge closure\ A =$
$UNIV$
⟨*proof*⟩

instance *metric-completion*::({*metric-space*, *second-countable-topology*}) *second-countable-topology*
⟨*proof*⟩

instance *metric-completion*::({*metric-space*, *second-countable-topology*}) *polish-space*
⟨*proof*⟩

end

8 Gromov hyperbolic spaces

theory *Gromov-Hyperbolicity*
imports *Isometries Metric-Completion*
begin

8.1 Definition, basic properties

Although we will mainly work with type classes later on, we introduce the definition of hyperbolicity on subsets of a metric space.

A set is δ-hyperbolic if it satisfies the following inequality. It is very obscure at first sight, but we will see several equivalent characterizations later on. For instance, a space is hyperbolic (maybe for a different constant δ) if all geodesic triangles are thin, i.e., every side is close to the union of the two other sides. This definition captures the main features of negative curvature at a large scale, and has proved extremely fruitful and influential.

Two important references on this topic are [GdlH90] and [BH99]. We will sometimes follow them, sometimes depart from them.

definition *Gromov-hyperbolic-subset*::*real* ⇒ (*'a*::*metric-space*) *set* ⇒ *bool*
where *Gromov-hyperbolic-subset delta A* = (∀ *x*∈*A*. ∀ *y*∈*A*. ∀ *z*∈*A*. ∀ *t*∈*A*. *dist x y* + *dist z t* ≤ *max* (*dist x z* + *dist y t*) (*dist x t* + *dist y z*) + *2* ∗ *delta*)

lemma *Gromov-hyperbolic-subsetI* [*intro*]:
assumes ⋀*x y z t*. *x* ∈ *A* ⟹ *y* ∈ *A* ⟹ *z* ∈ *A* ⟹ *t* ∈ *A* ⟹ *dist x y* + *dist z t* ≤ *max* (*dist x z* + *dist y t*) (*dist x t* + *dist y z*) + *2* ∗ *delta*
shows *Gromov-hyperbolic-subset delta A*
⟨*proof*⟩

When the four points are not all distinct, the above inequality is always satisfied for $\delta = 0$.

lemma *Gromov-hyperbolic-ineq-not-distinct*:
assumes *x* = *y* ∨ *x* = *z* ∨ *x* = *t* ∨ *y* = *z* ∨ *y* = *t* ∨ *z* = (*t*::*'a*::*metric-space*)
shows *dist x y* + *dist z t* ≤ *max* (*dist x z* + *dist y t*) (*dist x t* + *dist y z*)
⟨*proof*⟩

It readily follows from the definition that hyperbolicity passes to the closure of the set.

lemma *Gromov-hyperbolic-closure*:
assumes *Gromov-hyperbolic-subset delta A*
shows *Gromov-hyperbolic-subset delta* (*closure A*)

⟨*proof*⟩

A good formulation of hyperbolicity is in terms of Gromov products. Intuitively, the Gromov product of x and y based at e is the distance between e and the geodesic between x and y. It is also the time after which the geodesics from e to x and from e to y stop travelling together.

definition *Gromov-product-at*::(*'a*::*metric-space*) ⇒ *'a* ⇒ *'a* ⇒ *real*
  where *Gromov-product-at e x y* = (*dist e x* + *dist e y* − *dist x y*) / *2*

lemma *Gromov-hyperbolic-subsetI2*:
  fixes *delta*::*real*
  assumes ⋀*e x y z*. *e* ∈ *A* ⟹ *x* ∈ *A* ⟹ *y* ∈ *A* ⟹ *z* ∈ *A* ⟹ *Gromov-product-at* (*e*::*'a*::*metric-space*) *x z* ≥ *min* (*Gromov-product-at e x y*) (*Gromov-product-at e y z*) − *delta*
  shows *Gromov-hyperbolic-subset delta A*
⟨*proof*⟩

lemma *Gromov-product-nonneg* [*simp*, *mono-intros*]:
  Gromov-product-at e x y ≥ *0*
⟨*proof*⟩

lemma *Gromov-product-commute*:
  Gromov-product-at e x y = *Gromov-product-at e y x*
⟨*proof*⟩

lemma *Gromov-product-le-dist* [*simp*, *mono-intros*]:
  Gromov-product-at e x y ≤ *dist e x*
  Gromov-product-at e x y ≤ *dist e y*
⟨*proof*⟩

lemma *Gromov-product-le-infdist* [*mono-intros*]:
  assumes *geodesic-segment-between G x y*
  shows *Gromov-product-at e x y* ≤ *infdist e G*
⟨*proof*⟩

lemma *Gromov-product-add*:
  Gromov-product-at e x y + *Gromov-product-at x e y* = *dist e x*
⟨*proof*⟩

lemma *Gromov-product-geodesic-segment*:
  assumes *geodesic-segment-between G x y t* ∈ {*0*..*dist x y*}
  shows *Gromov-product-at x y* (*geodesic-segment-param G x t*) = *t*
⟨*proof*⟩

lemma *Gromov-product-e-x-x* [*simp*]:
  Gromov-product-at e x x = *dist e x*
⟨*proof*⟩

lemma *Gromov-product-at-diff*:

$|Gromov\text{-}product\text{-}at\ x\ y\ z - Gromov\text{-}product\text{-}at\ a\ b\ c| \leq dist\ x\ a + dist\ y\ b + dist\ z\ c$
⟨*proof*⟩

lemma *Gromov-product-at-diff1*:
$|Gromov\text{-}product\text{-}at\ a\ x\ y - Gromov\text{-}product\text{-}at\ b\ x\ y| \leq dist\ a\ b$
⟨*proof*⟩

lemma *Gromov-product-at-diff2*:
$|Gromov\text{-}product\text{-}at\ e\ x\ z - Gromov\text{-}product\text{-}at\ e\ y\ z| \leq dist\ x\ y$
⟨*proof*⟩

lemma *Gromov-product-at-diff3*:
$|Gromov\text{-}product\text{-}at\ e\ x\ y - Gromov\text{-}product\text{-}at\ e\ x\ z| \leq dist\ y\ z$
⟨*proof*⟩

The Gromov product is continuous in its three variables. We formulate it in terms of sequences, as it is the way it will be used below (and moreover continuity for functions of several variables is very poor in the library).

lemma *Gromov-product-at-continuous*:
assumes $(u \longrightarrow x)\ F\ (v \longrightarrow y)\ F\ (w \longrightarrow z)\ F$
shows $((\lambda n.\ Gromov\text{-}product\text{-}at\ (u\ n)\ (v\ n)\ (w\ n)) \longrightarrow Gromov\text{-}product\text{-}at\ x\ y\ z)\ F$
⟨*proof*⟩

8.2 Typeclass for Gromov hyperbolic spaces

We could (should?) just derive `Gromov_hyperbolic_space` from `metric_space`. However, in this case, properties of metric spaces are not available when working in the locale! It is more efficient to ensure that we have a metric space by putting a type class restriction in the definition. The δ in Gromov-hyperbolicity type class is called `deltaG` to avoid name clashes.

class *metric-space-with-deltaG* = *metric-space* +
fixes $deltaG::('a::metric\text{-}space)\ itself \Rightarrow real$

class *Gromov-hyperbolic-space* = *metric-space-with-deltaG* +
assumes *hyperb-quad-ineq0*: $Gromov\text{-}hyperbolic\text{-}subset\ (deltaG(TYPE('a::metric\text{-}space)))\ (UNIV::'a\ set)$

class *Gromov-hyperbolic-space-geodesic* = *Gromov-hyperbolic-space* + *geodesic-space*

lemma (**in** *Gromov-hyperbolic-space*) *hyperb-quad-ineq* [*mono-intros*]:
shows $dist\ x\ y + dist\ z\ t \leq max\ (dist\ x\ z + dist\ y\ t)\ (dist\ x\ t + dist\ y\ z) + 2 * deltaG(TYPE('a))$
⟨*proof*⟩

It readily follows from the definition that the completion of a δ-hyperbolic space is still δ-hyperbolic.

instantiation *metric-completion* :: (*Gromov-hyperbolic-space*) *Gromov-hyperbolic-space*
begin
definition *deltaG-metric-completion*::(*'a metric-completion*) *itself* ⇒ *real* **where**
deltaG-metric-completion - = deltaG(TYPE('a))

instance ⟨*proof*⟩
end

context *Gromov-hyperbolic-space*
begin

lemma *delta-nonneg* [*simp*, *mono-intros*]:
deltaG(TYPE('a)) ≥ 0
⟨*proof*⟩

lemma *hyperb-ineq* [*mono-intros*]:
Gromov-product-at (*e*::*'a*) *x z* ≥ *min* (*Gromov-product-at e x y*) (*Gromov-product-at e y z*) − *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

lemma *hyperb-ineq'* [*mono-intros*]:
Gromov-product-at (*e*::*'a*) *x z* + *deltaG*(*TYPE*(*'a*)) ≥ *min* (*Gromov-product-at e x y*) (*Gromov-product-at e y z*)
⟨*proof*⟩

lemma *hyperb-ineq-4-points* [*mono-intros*]:
Min {*Gromov-product-at* (*e*::*'a*) *x y*, *Gromov-product-at e y z*, *Gromov-product-at e z t*} − *2* ∗ *deltaG*(*TYPE*(*'a*)) ≤ *Gromov-product-at e x t*
⟨*proof*⟩

lemma *hyperb-ineq-4-points'* [*mono-intros*]:
Min {*Gromov-product-at* (*e*::*'a*) *x y*, *Gromov-product-at e y z*, *Gromov-product-at e z t*} ≤ *Gromov-product-at e x t* + *2* ∗ *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

In Gromov-hyperbolic spaces, geodesic triangles are thin, i.e., a point on one side of a geodesic triangle is close to the union of the two other sides (where the constant in "close" is 4δ, independent of the size of the triangle). We prove this basic property (which, in fact, is a characterization of Gromov-hyperbolic spaces: a geodesic space in which triangles are thin is hyperbolic).

lemma *thin-triangles1*:
assumes *geodesic-segment-between G x y geodesic-segment-between H x* (*z*::*'a*)
t ∈ {*0..Gromov-product-at x y z*}
shows *dist* (*geodesic-segment-param G x t*) (*geodesic-segment-param H x t*) ≤ *4* ∗ *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

theorem *thin-triangles*:

assumes *geodesic-segment-between Gxy x y*
geodesic-segment-between Gxz x z
geodesic-segment-between Gyz y z
$(w::'a) \in Gyz$
shows *infdist w* $(Gxy \cup Gxz) \leq 4 * deltaG(TYPE('a))$
⟨*proof*⟩

A consequence of the thin triangles property is that, although the geodesic between two points is in general not unique in a Gromov-hyperbolic space, two such geodesics are within $O(\delta)$ of each other.

lemma *geodesics-nearby*:
assumes *geodesic-segment-between G x y geodesic-segment-between H x y*
$(z::'a) \in G$
shows *infdist z H* $\leq 4 * deltaG(TYPE('a))$
⟨*proof*⟩

A small variant of the property of thin triangles is that triangles are slim, i.e., there is a point which is close to the three sides of the triangle (a "center" of the triangle, but only defined up to $O(\delta)$). And one can take it on any side, and its distance to the corresponding vertices is expressed in terms of a Gromov product.

lemma *slim-triangle*:
assumes *geodesic-segment-between Gxy x y*
geodesic-segment-between Gxz x z
geodesic-segment-between Gyz y $(z::'a)$
shows $\exists w.$ *infdist w Gxy* $\leq 4 * deltaG(TYPE('a)) \wedge$
infdist w Gxz $\leq 4 * deltaG(TYPE('a)) \wedge$
infdist w Gyz $\leq 4 * deltaG(TYPE('a)) \wedge$
dist w x $=$ (*Gromov-product-at x y z*) $\wedge\ w \in Gxy$
⟨*proof*⟩

The distance of a vertex of a triangle to the opposite side is essentially given by the Gromov product, up to 2δ.

lemma *dist-triangle-side-middle*:
assumes *geodesic-segment-between G x* $(y::'a)$
shows *dist z* (*geodesic-segment-param G x* (*Gromov-product-at x z y*)) $\leq$ *Gromov-product-at z x y* $+ 2 * deltaG(TYPE('a))$
⟨*proof*⟩

lemma *infdist-triangle-side* [*mono-intros*]:
assumes *geodesic-segment-between G x* $(y::'a)$
shows *infdist z G* $\leq$ *Gromov-product-at z x y* $+ 2 * deltaG(TYPE('a))$
⟨*proof*⟩

The distance of a point on a side of triangle to the opposite vertex is controlled by the length of the opposite sides, up to δ.

lemma *dist-le-max-dist-triangle*:

assumes *geodesic-segment-between G x y*
m ∈ *G*
shows *dist m z* ≤ *max* (*dist x z*) (*dist y z*) + *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

end

A useful variation around the previous properties is that quadrilaterals are thin, in the following sense: if one has a union of three geodesics from x to t, then a geodesic from x to t remains within distance 8δ of the union of these 3 geodesics. We formulate the statement in geodesic hyperbolic spaces as the proof requires the construction of an additional geodesic, but in fact the statement is true without this assumption, thanks to the Bonk-Schramm extension theorem.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *thin-quadrilaterals*:
assumes *geodesic-segment-between Gxy x y*
geodesic-segment-between Gyz y z
geodesic-segment-between Gzt z t
geodesic-segment-between Gxt x t
(*w*::'*a*) ∈ *Gxt*
shows *infdist w* (*Gxy* ∪ *Gyz* ∪ *Gzt*) ≤ *8* ∗ *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

There are converses to the above statements: if triangles are thin, or slim, then the space is Gromov-hyperbolic, for some δ. We prove these criteria here, following the proofs in Ghys (with a simplification in the case of slim triangles.

The basic result we will use twice below is the following: if points on sides of triangles at the same distance of the basepoint are close to each other up to the Gromov product, then the space is hyperbolic. The proof goes as follows. One wants to show that $(x, z)_e \geq \min((x, y)_e, (y, z)_e) - \delta = t - \delta$. On $[ex]$, $[ey]$ and $[ez]$, consider points wx, wy and wz at distance t of e. Then wx and wy are δ-close by assumption, and so are wy and wz. Then wx and wz are 2δ-close. One can use these two points to express $(x, z)_e$, and the result follows readily.

lemma (**in** *geodesic-space*) *controlled-thin-triangles-implies-hyperbolic*:
assumes ⋀(*x*::'*a*) *y z t Gxy Gxz*. *geodesic-segment-between Gxy x y* ⟹ *geodesic-segment-between Gxz x z* ⟹ *t* ∈ {*0*..*Gromov-product-at x y z*}
⟹ *dist* (*geodesic-segment-param Gxy x t*) (*geodesic-segment-param Gxz x t*) ≤ *delta*
shows *Gromov-hyperbolic-subset delta* (*UNIV*::'*a set*)
⟨*proof*⟩

We prove that if triangles are thin, i.e., they satisfy the Rips condition, i.e., every side of a triangle is included in the δ-neighborhood of the union of the other triangles, then the space is hyperbolic. If a point w on $[xy]$

satisfies $d(x,w) < (y,z)_x - \delta$, then its friend on $[xz] \cup [yz]$ has to be on $[xz]$, and roughly at the same distance of the origin. Then it follows that the point on $[xz]$ with $d(x,w') = d(x,w)$ is close to w, as desired. If $d(x,w) \in [(y,z)_x - \delta, (y,z)_x)$, we argue in the same way but for the point which is closer to x by an amount δ. Finally, the last case $d(x,w) = (y,z)_x$ follows by continuity.

proposition (**in** *geodesic-space*) *thin-triangles-implies-hyperbolic*:
assumes $\bigwedge(x::'a)$ *y z w Gxy Gyz Gxz. geodesic-segment-between Gxy x y* $\Longrightarrow$ *geodesic-segment-between Gxz x z* $\Longrightarrow$ *geodesic-segment-between Gyz y z*
$\Longrightarrow$ *w* $\in$ *Gxy* $\Longrightarrow$ *infdist w* (*Gxz* $\cup$ *Gyz*) $\leq$ *delta*
shows *Gromov-hyperbolic-subset* (*4* $*$ *delta*) (*UNIV*::*'a set*)
⟨*proof*⟩

Then, we prove that if triangles are slim (i.e., there is a point that is δ-close to all sides), then the space is hyperbolic. Using the previous statement, we should show that points on $[xy]$ and $[xz]$ at the same distance t of the origin are close, if $t \leq (y,z)_x$. There are two steps: - for $t = (y,z)_x$, then the two points are in fact close to the middle of the triangle (as this point satisfies $d(x,y) = d(x,w) + d(w,y) + O(\delta)$, and similarly for the other sides, one gets readily $d(x,w) = (y,z)_w + O(\delta)$ by expanding the formula for the Gromov product). Hence, they are close together. - For $t < (y,z)_x$, we argue that there are points $y' \in [xy]$ and $z' \in [xz]$ for which $t = (y',z')_x$, by a continuity argument and the intermediate value theorem. Then the result follows from the first step in the triangle $xy'z'$.

The proof we give is simpler than the one in [GdlH90], and gives better constants.

proposition (**in** *geodesic-space*) *slim-triangles-implies-hyperbolic*:
assumes $\bigwedge(x::'a)$ *y z Gxy Gyz Gxz. geodesic-segment-between Gxy x y* $\Longrightarrow$ *geodesic-segment-between Gxz x z* $\Longrightarrow$ *geodesic-segment-between Gyz y z*
$\Longrightarrow$ $\exists$ *w. infdist w Gxy* $\leq$ *delta* $\wedge$ *infdist w Gxz* $\leq$ *delta* $\wedge$ *infdist w Gyz* $\leq$ *delta*
shows *Gromov-hyperbolic-subset* (*6* $*$ *delta*) (*UNIV*::*'a set*)
⟨*proof*⟩

9 Metric trees

Metric trees have several equivalent definitions. The simplest one is probably that it is a geodesic space in which the union of two geodesic segments intersecting only at one endpoint is still a geodesic segment.

Metric trees are Gromov hyperbolic, with $\delta = 0$.

class *metric-tree* = *geodesic-space* +
assumes *geod-union*: *geodesic-segment-between G x y* $\Longrightarrow$ *geodesic-segment-between H y z* $\Longrightarrow$ *G* $\cap$ *H* = $\{y\}$ $\Longrightarrow$ *geodesic-segment-between* (*G* $\cup$ *H*) *x z*

We will now show that the real line is a metric tree, by identifying its geodesic segments, i.e., the compact intervals.

lemma *geodesic-segment-between-real*:
  assumes $x \leq (y::real)$
  shows *geodesic-segment-between* $(G::real\ set)\ x\ y = (G = \{x..y\})$
⟨*proof*⟩

lemma *geodesic-segment-between-real'*:
  $\{x{-}{-}y\} = \{min\ x\ y..max\ x\ (y::real)\}$
⟨*proof*⟩

lemma *geodesic-segment-real*:
  geodesic-segment $(G::real\ set) = (\exists\, x\ y.\ x \leq y \wedge G = \{x..y\})$
⟨*proof*⟩

instance *real::metric-tree*
⟨*proof*⟩

context *metric-tree* **begin**

We show that a metric tree is uniquely geodesic.

subclass *uniquely-geodesic-space*
⟨*proof*⟩

An important property of metric trees is that any geodesic triangle is degenerate, i.e., the three sides intersect at a unique point, the center of the triangle, that we introduce now.

definition $center::'a \Rightarrow 'a \Rightarrow 'a \Rightarrow 'a$
  where $center\ x\ y\ z = (SOME\ t.\ t \in \{x{-}{-}y\} \cap \{x{-}{-}z\} \cap \{y{-}{-}z\})$

lemma *center-as-intersection*:
  $\{x{-}{-}y\} \cap \{x{-}{-}z\} \cap \{y{-}{-}z\} = \{center\ x\ y\ z\}$
⟨*proof*⟩

lemma *center-on-geodesic* [*simp*]:
  $center\ x\ y\ z \in \{x{-}{-}y\}$
  $center\ x\ y\ z \in \{x{-}{-}z\}$
  $center\ x\ y\ z \in \{y{-}{-}z\}$
  $center\ x\ y\ z \in \{y{-}{-}x\}$
  $center\ x\ y\ z \in \{z{-}{-}x\}$
  $center\ x\ y\ z \in \{z{-}{-}y\}$
⟨*proof*⟩

lemma *center-commute*:
  $center\ x\ y\ z = center\ x\ z\ y$
  $center\ x\ y\ z = center\ y\ x\ z$
  $center\ x\ y\ z = center\ y\ z\ x$
  $center\ x\ y\ z = center\ z\ x\ y$

center x y z = center z y x
⟨*proof*⟩

lemma *center-dist*:
dist x (center x y z) = Gromov-product-at x y z
⟨*proof*⟩

lemma *geodesic-intersection*:
$\{x{-}{-}y\} \cap \{x{-}{-}z\} = \{x{-}{-}center\ x\ y\ z\}$
⟨*proof*⟩
end

We can now prove that a metric tree is Gromov hyperbolic, for $\delta = 0$. The simplest proof goes through the slim triangles property: it suffices to show that, given a geodesic triangle, there is a point at distance at most 0 of each of its sides. This is the center we have constructed above.

class *metric-tree-with-delta = metric-tree + metric-space-with-deltaG +*
assumes *delta0*: *deltaG(TYPE('a::metric-space)) = 0*

class *Gromov-hyperbolic-space-0 = Gromov-hyperbolic-space +*
assumes *delta0* [*simp*]: *deltaG(TYPE('a::metric-space)) = 0*

class *Gromov-hyperbolic-space-0-geodesic = Gromov-hyperbolic-space-0 + geodesic-space*

Isabelle does not accept cycles in the class graph. So, we will show that `metric_tree_with_delta` is a subclass of `Gromov_hyperbolic_space_0_geodesic`, and conversely that `Gromov_hyperbolic_space_0_geodesic` is a subclass of `metric_tree`.

In a tree, we have already proved that triangles are 0-slim (the center is common to all sides of the triangle). The 0-hyperbolicity follows from one of the equivalent characterizations of hyperbolicity (the other characterizations could be used as well, but the proofs would be less immediate.)

subclass (**in** *metric-tree-with-delta*) *Gromov-hyperbolic-space-0*
⟨*proof*⟩

To use the fact that reals are Gromov hyperbolic, given that they are a metric tree, we need to instantiate them as `metric_tree_with_delta`.

instantiation *real::metric-tree-with-delta*
begin
definition *deltaG-real::real itself ⇒ real*
where *deltaG-real - = 0*
instance ⟨*proof*⟩
end

Let us now prove the converse: a geodesic space which is δ-hyperbolic for $\delta = 0$ is a metric tree. For the proof, we consider two geodesic segments $G = [x, y]$ and $H = [y, z]$ with a common endpoint, and we have to show that their union is still a geodesic segment from x to z. For this, introduce a geodesic segment $L = [x, z]$. By the property of thin triangles, G is included in $H \cup L$. In particular, a point Y close to y but different from y on G is on L, and therefore realizes the equality $d(x, z) = d(x, Y) + d(Y, z)$. Passing to the limit, y also satisfies this equality. The conclusion readily follows thanks to Lemma `geodesic_segment_union`.

subclass (**in** *Gromov-hyperbolic-space-0-geodesic*) *metric-tree*
⟨*proof*⟩

end

theory *Morse-Gromov-Theorem*
imports *HOL−Decision-Procs.Approximation Gromov-Hyperbolicity Hausdorff-Distance*
begin

hide-const (**open**) *Approximation.Min*
hide-const (**open**) *Approximation.Max*

10 Quasiconvexity

In a Gromov-hyperbolic setting, convexity is not a well-defined notion as everything should be coarse. The good replacement is quasi-convexity: A set X is C-quasi-convex if any pair of points in X can be joined by a geodesic that remains within distance C of X. One could also require this for all geodesics, up to changing C, as two geodesics between the same endpoints remain within uniformly bounded distance. We use the first definition to ensure that a geodesic is 0-quasi-convex.

definition *quasiconvex*::*real* ⇒ (*'a*::*metric-space*) *set* ⇒ *bool*
where *quasiconvex C X* = (*C* ≥ *0* ∧ (∀ *x*∈*X*. ∀ *y*∈*X*. ∃ *G*. *geodesic-segment-between G x y* ∧ (∀ *z*∈*G*. *infdist z X* ≤ *C*)))

lemma *quasiconvexD*:
assumes *quasiconvex C X x* ∈ *X y* ∈ *X*
shows ∃ *G*. *geodesic-segment-between G x y* ∧ (∀ *z*∈*G*. *infdist z X* ≤ *C*)
⟨*proof*⟩

lemma *quasiconvexC*:
assumes *quasiconvex C X*
shows *C* ≥ *0*
⟨*proof*⟩

lemma *quasiconvexI*:
assumes *C* ≥ *0*
⋀*x y*. *x* ∈ *X* ⟹ *y* ∈ *X* ⟹ (∃ *G*. *geodesic-segment-between G x y* ∧

$(\forall z \in G.\ infdist\ z\ X \leq C))$
shows *quasiconvex C X*
⟨*proof*⟩

lemma *quasiconvex-of-geodesic*:
assumes *geodesic-segment G*
shows *quasiconvex 0 G*
⟨*proof*⟩

lemma *quasiconvex-empty*:
assumes $C \geq 0$
shows *quasiconvex C* {}
⟨*proof*⟩

lemma *quasiconvex-mono*:
assumes $C \leq D$
quasiconvex C G
shows *quasiconvex D G*
⟨*proof*⟩

The r-neighborhood of a quasi-convex set is still quasi-convex in a hyperbolic space, for a constant that does not depend on r.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *quasiconvex-thickening*:
assumes *quasiconvex C* $(X::'a\ set)$ $r \geq 0$
shows *quasiconvex* $(C + 8 * deltaG(TYPE('a)))$ $(\bigcup x \in X.\ cball\ x\ r)$
⟨*proof*⟩

If x has a projection p on a quasi-convex set G, then all segments from a point in G to x go close to p, i.e., the triangular inequality $d(x,y) \leq d(x,p)+d(p,y)$ is essentially an equality, up to an additive constant.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *dist-along-quasiconvex*:
assumes *quasiconvex C G* $p \in$ *proj-set x G* $y \in G$
shows $dist\ x\ p + dist\ p\ y \leq dist\ x\ y + 4 * deltaG(TYPE('a)) + 2 * C$
⟨*proof*⟩

The next lemma is [CDP90, Proposition 10.2.1] with better constants. It states that the distance between the projections on a quasi-convex set is controlled by the distance of the original points, with a gain given by the distances of the points to the set.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *proj-along-quasiconvex-contraction*:
assumes *quasiconvex C G* $px \in$ *proj-set x G* $py \in$ *proj-set y G*
shows $dist\ px\ py \leq max\ (5 * deltaG(TYPE('a)) + 2 * C)\ (dist\ x\ y - dist\ px\ x - dist\ py\ y + 10 * deltaG(TYPE('a)) + 4 * C)$
⟨*proof*⟩

The projection on a quasi-convex set is 1-Lipschitz up to an additive error.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *proj-along-quasiconvex-contraction'*:
assumes *quasiconvex C G* $px \in$ *proj-set x G* $py \in$ *proj-set y G*

shows *dist px py* ≤ *dist x y* + *4* ∗ *deltaG*(*TYPE*(*'a*)) + *2* ∗ *C*
⟨*proof*⟩

We can in particular specialize the previous statements to geodesics, which are 0-quasi-convex.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *dist-along-geodesic*:
assumes *geodesic-segment G p* ∈ *proj-set x G y* ∈ *G*
shows *dist x p* + *dist p y* ≤ *dist x y* + *4* ∗ *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *proj-along-geodesic-contraction*:
assumes *geodesic-segment G px* ∈ *proj-set x G py* ∈ *proj-set y G*
shows *dist px py* ≤ *max* (*5* ∗ *deltaG*(*TYPE*(*'a*))) (*dist x y* − *dist px x* − *dist py y* + *10* ∗ *deltaG*(*TYPE*(*'a*)))
⟨*proof*⟩

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *proj-along-geodesic-contraction'*:
assumes *geodesic-segment G px* ∈ *proj-set x G py* ∈ *proj-set y G*
shows *dist px py* ≤ *dist x y* + *4* ∗ *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

If one projects a continuous curve on a quasi-convex set, the image does not have to be connected (the projection is discontinuous), but since the projections of nearby points are within uniformly bounded distance one can find in the projection a point with almost prescribed distance to the starting point, say. For further applications, we also pick the first such point, i.e., all the previous points are also close to the starting point.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *quasi-convex-projection-small-gaps*:
assumes *continuous-on* {*a*..(*b*::*real*)} *f*
a ≤ *b*
quasiconvex C G
⋀*t*. *t* ∈ {*a*..*b*} ⟹ *p t* ∈ *proj-set* (*f t*) *G*
delta > *deltaG*(*TYPE*(*'a*))
d ∈ {*4* ∗ *delta* + *2* ∗ *C*..*dist* (*p a*) (*p b*)}
shows ∃ *t* ∈ {*a*..*b*}. (*dist* (*p a*) (*p t*) ∈ {*d* − *4* ∗ *delta* − *2* ∗ *C* .. *d*})
∧ (∀ *s* ∈ {*a*..*t*}. *dist* (*p a*) (*p s*) ≤ *d*)
⟨*proof*⟩

Same lemma, except that one exchanges the roles of the beginning and the end point.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *quasi-convex-projection-small-gaps'*:
assumes *continuous-on* {*a*..(*b*::*real*)} *f*
a ≤ *b*
quasiconvex C G
⋀*x*. *x* ∈ {*a*..*b*} ⟹ *p x* ∈ *proj-set* (*f x*) *G*
delta > *deltaG*(*TYPE*(*'a*))
d ∈ {*4* ∗ *delta* + *2* ∗ *C*..*dist* (*p a*) (*p b*)}
shows ∃ *t* ∈ {*a*..*b*}. *dist* (*p b*) (*p t*) ∈ {*d* − *4* ∗ *delta* − *2* ∗ *C* .. *d*}

$\wedge\ (\forall s \in \{t..b\}.\ dist\ (p\ b)\ (p\ s) \leq d)$

⟨*proof*⟩

11 The Morse-Gromov Theorem

The goal of this section is to prove a central basic result in the theory of hyperbolic spaces, usually called the Morse Lemma. It is really a theorem, and we add the name Gromov the avoid the confusion with the other Morse lemma on the existence of good coordinates for C^2 functions with non-vanishing hessian.

It states that a quasi-geodesic remains within bounded distance of a geodesic with the same endpoints, the error depending only on δ and on the parameters (λ, C) of the quasi-geodesic, but not on its length.

There are several proofs of this result. We will follow the one of Shchur [Shc13], which gets an optimal dependency in terms of the parameters of the quasi-isometry, contrary to all previous proofs. The price to pay is that the proof is more involved (relying in particular on the fact that the closest point projection on quasi-convex sets is exponentially contracting).

We will also give afterwards for completeness the proof in [BH99], as it brings up interesting tools, although the dependency it gives is worse.

The next lemma (for $C = 0$, Lemma 2 in [Shc13]) asserts that, if two points are not too far apart (at distance at most 10δ), and far enough from a given geodesic segment, then when one moves towards this geodesic segment by a fixed amount (here 5δ), then the two points become closer (the new distance is at most 5δ, gaining a factor of 2). Later, we will iterate this lemma to show that the projection on a geodesic segment is exponentially contracting. For the application, we give a more general version involving an additional constant C.

This lemma holds for δ the hyperbolicity constant. We will want to apply it with $\delta > 0$, so to avoid problems in the case $\delta = 0$ we formulate it not using the hyperbolicity constant of the given type, but any constant which is at least the hyperbolicity constant (this is to work around the fact that one can not say or use easily in Isabelle that a type with hyperbolicity δ is also hyperbolic for any larger constant δ'.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *geodesic-projection-exp-contracting-aux*:
 assumes *geodesic-segment G*
 px ∈ *proj-set x G*
 py ∈ *proj-set y G*
 delta ≥ *deltaG*(*TYPE*(*'a*))
 dist x y ≤ *10* ∗ *delta* + *C*
 M ≥ *15/2* ∗ *delta*
 dist px x ≥ *M* + *5* ∗ *delta* + *C/2*
 dist py y ≥ *M* + *5* ∗ *delta* + *C/2*

$C \geq 0$
shows *dist* (*geodesic-segment-param* {*px*−−*x*} *px M*)
(*geodesic-segment-param* {*py*−−*y*} *py M*) ≤ *5* ∗ *delta*
⟨*proof*⟩

The next lemma (Lemma 10 in [Shc13] for $C = 0$) asserts that the projection on a geodesic segment is an exponential contraction. More precisely, if a path of length L is at distance at least D of a geodesic segment G, then the projection of the path on G has diameter at most $CL\exp(-cD/\delta)$, where C and c are universal constants. This is not completely true at one can not go below a fixed size, as always, so the correct bound is $K\max(\delta, L\exp(-cD/\delta))$. For the application, we give a slightly more general statement involving an additional constant C.

This statement follows from the previous lemma: if one moves towards G by 10δ, then the distance between points is divided by 2. Then one iterates this statement as many times as possible, gaining a factor 2 each time and therefore an exponential factor in the end.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *geodesic-projection-exp-contracting*:
assumes *geodesic-segment G*
⋀*x y*. *x* ∈ {*a*..*b*} ⟹ *y* ∈ {*a*..*b*} ⟹ *dist* (*f x*) (*f y*) ≤ *lambda* ∗ *dist x y* + *C*
a ≤ *b*
pa ∈ *proj-set* (*f a*) *G*
pb ∈ *proj-set* (*f b*) *G*
⋀*t*. *t* ∈ {*a*..*b*} ⟹ *infdist* (*f t*) *G* ≥ *D*
D ≥ *15*/*2* ∗ *delta* + *C*/*2*
delta > *deltaG*(*TYPE*('*a*))
C ≥ *0*
lambda ≥ *0*
shows *dist pa pb* ≤ *max* (*5* ∗ *deltaG*(*TYPE*('*a*))) ((*4* ∗ *exp*(*1*/*2* ∗ *ln 2*)) ∗ *lambda* ∗ (*b*−*a*) ∗ *exp*(−(*D*−*C*/*2*) ∗ *ln 2* / (*5* ∗ *delta*)))
⟨*proof*⟩

We deduce from the previous result that a projection on a quasiconvex set is also exponentially contracting. To do this, one uses the contraction of a projection on a geodesic, and one adds up the additional errors due to the quasi-convexity. In particular, the projections on the original quasiconvex set or the geodesic do not have to coincide, but they are within distance at most $C + 8\delta$.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *quasiconvex-projection-exp-contracting*:
assumes *quasiconvex K G*
⋀*x y*. *x* ∈ {*a*..*b*} ⟹ *y* ∈ {*a*..*b*} ⟹ *dist* (*f x*) (*f y*) ≤ *lambda* ∗ *dist x y* + *C*
a ≤ *b*
pa ∈ *proj-set* (*f a*) *G*
pb ∈ *proj-set* (*f b*) *G*
⋀*t*. *t* ∈ {*a*..*b*} ⟹ *infdist* (*f t*) *G* ≥ *D*

$D \geq 15/2 * delta + K + C/2$
$delta > deltaG(TYPE('a))$
$C \geq 0$
$lambda \geq 0$
shows $dist\ pa\ pb \leq 2 * K + 8 * delta + max\ (5 * deltaG(TYPE('a)))\ ((4 * exp(1/2 * ln\ 2)) * lambda * (b-a) * exp(-(D - K - C/2) * ln\ 2\ /\ (5 * delta)))$
$\langle proof \rangle$

The next statement is the main step in the proof of the Morse-Gromov theorem given by Shchur in [Shc13], asserting that a quasi-geodesic and a geodesic with the same endpoints are close. We show that a point on the quasi-geodesic is close to the geodesic – the other inequality will follow easily later on. We also assume that the quasi-geodesic is parameterized by a Lipschitz map – the general case will follow as any quasi-geodesic can be approximated by a Lipschitz map with good controls.

Here is a sketch of the proof. Fix two large constants $L \leq D$ (that we will choose carefully to optimize the values of the constants at the end of the proof). Consider a quasi-geodesic f between two points $f(u^-)$ and $f(u^+)$, and a geodesic segment G between the same points. Fix $f(z)$. We want to find a bound on $d(f(z), G)$. 1 - If this distance is smaller than L, we are done (and the bound is L). 2 - Assume it is larger. Let π_z be a projection of $f(z)$ on G (at distance $d(f(z), G)$ of $f(z)$), and H a geodesic between z and π_z. The idea will be to project the image of f on H, and record how much progress is made towards $f(z)$. In this proof, we will construct several points before and after z. When necessary, we put an exponent $-$ on the points before z, and $+$ on the points after z. To ease the reading, the points are ordered following the alphabetical order, i.e., $u^- \leq v \leq w \leq x \leq y^- \leq z$.

One can find two points $f(y^-)$ and $f(y^+)$ on the left and the right of $f(z)$ that project on H roughly at distance L of π_z (up to some $O(\delta)$ – recall that the closest point projection is not uniquely defined, and not continuous, so we make some choice here). Let d^- be the minimal distance of $f([u^-, y^-])$ to H, and let d^+ be the minimal distance of $f([y^+, u^+)]$ to H.

2.1 If the two distances d^- and d^+ are less than D, then the distance between two points realizing the minimum (say $f(c^-)$ and $f(c^+)$) is at most $2D + L$, hence $c^+ - c^-$ is controlled (by $\lambda \cdot (2D + L) + C$) thanks to the quasi-isometry property. Therefore, $f(z)$ is not far away from $f(c^-)$ and $f(c^+)$ (again by the quasi-isometry property). Since the distance from these points to π_z is controlled (by $D + L$), we get a good control on $d(f(z), \pi_z)$, as desired.

2.2 The interesting case is when d^- and d^+ are both $> D$. Assume also for instance $d^- \geq d^+$, as the other case is analogous. We will construct two points $f(v)$ and $f(x)$ with $u^- \leq v \leq x \leq y^-$ with the following property:

$$K_1 e^{K_2 d(f(v), H)} \leq x - v, \tag{1}$$

where K_1 and K_2 are some explicit constants (depending on λ, δ, L and D).

Let us show how this will conclude the proof. The distance from $f(v)$ to $f(c^+)$ is at most $d(f(v), H) + L + d^+ \leq 3d(f(v), H)$. Therefore, $c^+ - v$ is also controlled by $K'd(f(v), H)$ by the quasi-isometry property. This gives

$$\begin{aligned} K &\leq K(x-v)e^{-K(c^+-v)} \leq (e^{K(x-v)} - 1) \cdot e^{-K(c^+-v)} \\ &= e^{-K(c^+-x)} - e^{-K(c^+-v)} \leq e^{-K(c^+-x)} - e^{-K(u^+-u^-)}. \end{aligned}$$

This shows that, when one goes from the original quasi-geodesic $f([u^-, u^+])$ to the restricted quasi-geodesic $f([x, c^+])$, the quantity $e^{-K\cdot}$ decreases by a fixed amount. In particular, this process can only happen a uniformly bounded number of times, say n.

Let G' be a geodesic between $f(x)$ and $f(c^+)$. One checks geometrically that $d(f(z), G) \leq d(f(z), G') + (L + O(\delta))$, as both projections of $f(x)$ and $f(c^+)$ on H are within distance L of π_z. Iterating the process n times, one gets finally $d(f(z), G) \leq O(1) + n(L + O(\delta))$. This is the desired bound for $d(f(z), G)$.

To complete the proof, it remains to construct the points $f(v)$ and $f(x)$ satisfying (1). This will be done through an inductive process.

Assume first that there is a point $f(v)$ whose projection on H is close to the projection of $f(u^-)$, and with $d(f(v), H) \leq 2d^-$. Then the projections of $f(v)$ and $f(y^-)$ are far away (at distance at least $L + O(\delta)$). Since the portion of f between v and y^- is everywhere at distance at least d^- of H, the projection on H contracts by a factor e^{-d^-}. It follows that this portion of f has length at least $e^{d^-} \cdot (L + O(\delta))$. Therefore, by the quasi-isometry property, one gets $x - v \geq Ke^{d^-}$. On the other hand, $d(v, H)$ is bounded above by $2d^-$ by assumption. This gives the desired inequality (1) with $x = y^-$.

Otherwise, all points $f(v)$ whose projection on H is close to the projection of $f(u^-)$ are such that $d(f(v), H) \geq 2d^-$. Consider $f(w_1)$ a point whose projection on H is at distance roughly 10δ of the projection of $f(u^-)$. Let V_1 be the set of points at distance at most d^- of H, i.e., the d_1-neighborhood of H. Then the distance between the projections of $f(u^-)$ and $f(w_1)$ on V_1 is very large (are there is an additional big contraction to go from V_1 to H). And moreover all the intermediate points $f(v)$ are at distance at least $2d^-$ of H, and therefore at distance at least d^- of H. Then one can play the same game as in the first case, where y^- replaced by w_1 and H replaced by V_1. If there is a point $f(v)$ whose projection on V_1 is close to the projection of $f(u^-)$, then the pair of points v and $x = w_1$ works. Otherwise, one lifts everything to V_2, the neighborhood of size $2d^-$ of V_1, and one argues again in the same way.

The induction goes on like this until one finds a suitable pair of points. The process has indeed to stop at one time, as it can only go on while $f(u^-)$ is outside of V_k, the $(2^k - 1)d^-$ neighborhood of H). This concludes the sketch of the proof, modulo the adjustment of constants.

Comments on the formalization below:

- The proof is written as an induction on $u^+ - u^-$. This makes it possible to either prove the bound directly (in the cases 1 and 2.1 above), or to use the bound on $d(z, G')$ in case 2.2 using the induction assumption, and conclude the proof. Of course, $u^+ - u^-$ is not integer-valued, but in the reduction to G' it decays by a fixed amount, so one can easily write this down as a genuine induction.

- The main difficulty in the proof is to construct the pair (v, x) in case 2.2. This is again written as an induction over k: either the required bound is true, or one can find a point $f(w)$ whose projection on V_k is far enough from the projection of $f(u^-)$. Then, either one can use this point to prove the bound, or one can construct a point with the same property with respect to V_{k+1}, concluding the induction.

- Instead of writing u^- and u^+ (which are not good variable names in Isabelle), we write um and uM. Similarly for other variables.

- The proof only works when $\delta > 0$ (as one needs to divide by δ in the exponential gain). Hence, we formulate it for some δ which is strictly larger than the hyperbolicity constant. In a subsequent application of the lemma, we will deduce the same statement for the hyperbolicity constant by a limiting argument.

- To optimize the value of the constant in the end, there is an additional important trick with respect to the above sketch: in case 2.2, there is an exponential gain. One can spare a fraction $(1 - \alpha)$ of this gain to improve the constants, and spend the remaining fraction α to make the argument work. This makes it possible to reduce the value of the constant roughly from 40000 to 100, so we do it in the proof below. The values of L, D and α can be chosen freely, and have been chosen to get the best possible constant in the end.

- For another optimization, we do not induce in terms of the distance from $f(z)$ to the geodesic G, but rather in terms of the Gromov product $(f(u^-), f(u^+))_{f(z)}$ (which is the same up to $O(\delta)$. And we do not take for H a geodesic from $f(z)$ to its projection on G, but rather a geodesic from $f(z)$ to the point m on $[f(u^-), f(u^+)]$ opposite to $f(z)$ in the tripod, i.e., at distance $(f(z), f(u^+))_{f(u^-)}$ of $f(u^-)$, and at distance $(f(z), f(u^-))_{f(u^+)}$ of $f(u^+)$. Let π_z denote the point on $[f(z), m]$ at distance $(f(u^-), f(u^+)_{f(z)}$ of $f(z)$. (It is within distance 2δ of m). In both approaches, what we want to control by induction is the distance from $f(z)$ to π_z. However, in the first approach, the points $f(u^-)$ and $f(u^+)$ project on H between π_z and $f(z)$, and since the location of their projection is only controlled up to 4δ one loses essentially a 4δ-length of L for the forthcoming argument. In the second approach, the projections on H are on the other side of π_z compared to $f(z)$, so one does not lose anything, and in the end it gives genuinely better bounds (making it possible to gain roughly 10δ in the final estimate).

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *Morse-Gromov-theorem-aux1*:
fixes *f*::*real* ⇒ *'a*
assumes *continuous-on* {*a..b*} *f*
lambda C−quasi-isometry-on {*a..b*} *f*
a ≤ *b*
geodesic-segment-between G (*f a*) (*f b*)
z ∈ {*a..b*}
delta > *deltaG*(*TYPE*(*'a*))
shows *infdist* (*f z*) *G* ≤ *lambda*^*2* ∗ (*11/2* ∗ *C* + *91* ∗ *delta*)
⟨*proof*⟩

Still assuming that our quasi-isometry is Lipschitz, we will improve slightly on the previous result, first going down to the hyperbolicity constant of the space, and also showing that, conversely, the geodesic is contained in a neighborhood of the quasi-geodesic. The argument for this last point goes as follows. Consider a point x on the geodesic. Define two sets to be the D-thickenings of $[a, x]$ and $[x, b]$ respectively, where D is such that any point on the quasi-geodesic is within distance D of the geodesic (as given by the previous theorem). The union of these two sets covers the quasi-geodesic, and they are both closed and nonempty. By connectedness, there is a point z in their intersection, D-close both to a point x^- before x and to a point x^+ after x. Then x belongs to a geodesic between x^- and x^+, which is contained in a 4δ-neighborhood of geodesics from x^+ to z and from x^- to z by hyperbolicity. It follows that x is at distance at most $D + 4\delta$ of z, concluding the proof.

lemma (**in** *Gromov-hyperbolic-space-geodesic*) *Morse-Gromov-theorem-aux2*:
fixes *f*::*real* ⇒ *'a*
assumes *continuous-on* {*a..b*} *f*
lambda C−quasi-isometry-on {*a..b*} *f*
geodesic-segment-between G (*f a*) (*f b*)
shows *hausdorff-distance* (*f'*{*a..b*}) *G* ≤ *lambda*^*2* ∗ (*11/2* ∗ *C* + *92* ∗ *deltaG*(*TYPE*(*'a*)))
⟨*proof*⟩

The full statement of the Morse-Gromov Theorem, asserting that a quasi-geodesic is within controlled distance of a geodesic with the same endpoints. It is given in the formulation of Shchur [Shc13], with optimal control in terms of the parameters of the quasi-isometry. This statement follows readily from the previous one and from the fact that quasi-geodesics can be approximated by Lipschitz ones.

theorem (**in** *Gromov-hyperbolic-space-geodesic*) *Morse-Gromov-theorem*:

fixes *f::real* ⇒ *'a*
assumes *lambda C−quasi-isometry-on {a..b} f*
geodesic-segment-between G (f a) (f b)
shows *hausdorff-distance (f'{a..b}) G* ≤ *92 * lambda^2 * (C + deltaG(TYPE('a)))*
⟨*proof*⟩

This theorem implies the same statement for two quasi-geodesics sharing their endpoints.

theorem (**in** *Gromov-hyperbolic-space-geodesic*) *Morse-Gromov-theorem2*:
fixes *c d::real* ⇒ *'a*
assumes *lambda C−quasi-isometry-on {A..B} c*
lambda C−quasi-isometry-on {A..B} d
c A = d A c B = d B
shows *hausdorff-distance (c'{A..B}) (d'{A..B})* ≤ *184 * lambda^2 * (C + deltaG(TYPE('a)))*
⟨*proof*⟩

We deduce from the Morse lemma that hyperbolicity is invariant under quasi-isometry.

First, we note that the image of a geodesic segment under a quasi-isometry is close to a geodesic segment in Hausdorff distance, as it is a quasi-geodesic.

lemma *geodesic-quasi-isometric-image*:
fixes *f::'a::metric-space* ⇒ *'b::Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry-on UNIV f*
geodesic-segment-between G x y
shows *hausdorff-distance (f'G) {f x−−f y}* ≤ *92 * lambda^2 * (C + deltaG(TYPE('b)))*
⟨*proof*⟩

We deduce that hyperbolicity is invariant under quasi-isometry. The proof goes as follows: we want to see that a geodesic triangle is delta-thin, i.e., a point on a side Gxy is close to the union of the two other sides Gxz and Gyz. Pull everything back by the quasi-isometry: we obtain three quasi-geodesic, each of which is close to the corresponding geodesic segment by the Morse lemma. As the geodesic triangle is thin, it follows that the quasi-geodesic triangle is also thin, i.e., a point on $f^{-1}Gxy$ is close to $f^{-1}Gxz \cup f^{-1}Gyz$ (for some explicit, albeit large, constant). Then push everything forward by f: as it is a quasi-isometry, it will again distort distances by a bounded amount.

lemma *Gromov-hyperbolic-invariant-under-quasi-isometry-explicit*:
fixes *f::'a::geodesic-space* ⇒ *'b::Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
shows *Gromov-hyperbolic-subset (752 * lambda^3 * (C + deltaG(TYPE('b)))) (UNIV::('a set))*
⟨*proof*⟩

Most often, the precise value of the constant in the previous theorem is irrelevant, it is used in the following form.

theorem *Gromov-hyperbolic-invariant-under-quasi-isometry*:
assumes *quasi-isometric* (*UNIV*::(*'a*::*geodesic-space*) *set*) (*UNIV*::(*'b*::*Gromov-hyperbolic-space-geodesic*) *set*)
shows $\exists$ *delta*. *Gromov-hyperbolic-subset delta* (*UNIV*::*'a set*)
⟨*proof*⟩

A central feature of hyperbolic spaces is that a path from x to y can not deviate too much from a geodesic from x to y unless it is extremely long (exponentially long in terms of the distance from x to y). This is useful both to ensure that short paths (for instance quasi-geodesics) stay close to geodesics, see the Morse lemme below, and to ensure that paths that avoid a given large ball of radius R have to be exponentially long in terms of R (this is extremely useful for random walks). This proposition is the first non-trivial result on hyperbolic spaces in [BH99] (Proposition III.H.1.6). We follow their proof.

The proof is geometric, and uses the existence of geodesics and the fact that geodesic triangles are thin. In fact, the result still holds if the space is not geodesic, as it can be deduced by embedding the hyperbolic space in a geodesic hyperbolic space and using the result there.

proposition (**in** *Gromov-hyperbolic-space-geodesic*) *lipschitz-path-close-to-geodesic*:
fixes *c*::*real* $\Rightarrow$ *'a*
assumes *M*−*lipschitz-on* {*A*..*B*} *c*
geodesic-segment-between G (*c A*) (*c B*)
$x \in G$
shows *infdist x* (*c*‘{*A*..*B*}) $\leq$ (*4*/*ln 2*) $*$ *deltaG*(*TYPE*(*'a*)) $*$ *max 0* (*ln* (*B*−*A*)) + *M*
⟨*proof*⟩

By rescaling coordinates at the origin, one obtains a variation around the previous statement.

proposition (**in** *Gromov-hyperbolic-space-geodesic*) *lipschitz-path-close-to-geodesic'*:
fixes *c*::*real* $\Rightarrow$ *'a*
assumes *M*−*lipschitz-on* {*A*..*B*} *c*
geodesic-segment-between G (*c A*) (*c B*)
$x \in G$
a > *0*
shows *infdist x* (*c*‘{*A*..*B*}) $\leq$ (*4*/*ln 2*) $*$ *deltaG*(*TYPE*(*'a*)) $*$ *max 0* (*ln* (*a* $*$ (*B*−*A*))) + *M*/*a*
⟨*proof*⟩

We can now give another proof of the Morse-Gromov Theorem, as described in [BH99]. It is more direct than the one we have given above, but it gives a worse dependence in terms of the quasi-isometry constants. In particular, when $C = \delta = 0$, it does not recover the fact that a quasi-geodesic has to coincide with a geodesic.

theorem (**in** *Gromov-hyperbolic-space-geodesic*) *Morse-Gromov-theorem-BH-proof*:

fixes *c::real* ⇒ *'a*
assumes *lambda C−quasi-isometry-on {A..B} c*
shows *hausdorff-distance (c'{A..B}) {c A−−c B}* ≤ *72 ∗ lambda^2 ∗ (C + lambda + deltaG(TYPE('a))^2)*
⟨*proof*⟩

end

12 The Bonk Schramm extension

theory *Bonk-Schramm-Extension*
imports *Morse-Gromov-Theorem*
begin

We want to show that any metric space is isometrically embedded in a metric space which is geodesic (i.e., there is an embedded geodesic between any two points) and complete. There are many such constructions, but a very interesting one has been given by Bonk and Schramm in [BS00], together with an additional property of the completion: if the space is delta-hyperbolic (in the sense of Gromov), then its completion also is, with the same constant delta. It follows in particular that a 0-hyperbolic space embeds in a 0-hyperbolic geodesic space, i.e., a metric tree (there is an easier direct construction in this case).

Another embedding of a metric space in a geodesic one is constructed by Mineyev [Min05], it is more canonical in a sense (isometries of the original space extend to the new space), but it is not clear if it preserves hyperbolicity.

The argument of Bonk and Schramm goes as follows: - first, if one wants to add the middle of a pair of points a and b in a space E, there is a nice formula for the distance on a new space $E \cup \{*\}$ (where $*$ will by construction be a middle of a and b). - by transfinite induction on all the pair of points in the space, one adds all the missing middles - then one completes the space - then one adds all the middles - then one goes on like that, transfinitely many times - at some point, the process stops for cardinality reasons

The resulting space is complete and has middles for all pairs of points. It is then standard that it is geodesic (this is proved in `Geodesic_Spaces.thy`).

Implementing this construction in Isabelle is interesting and nontrivial, as transfinite induction is not that easy, especially when intermingled with metric completion (i.e., taking the quotient space of all Cauchy sequences). In particular, taking sequences of metric completions would mean changing types at each step, along a transfinite number of steps. It does not seem possible to do it naively in this way.

We avoid taking quotients in the middle of the argument, as this is too messy. Instead, we define a pseudo-distance (i.e., a function satisyfing the triangular inequality, but such that $d(x,y)$ can vanish even if x and y are different) on an increasing set, which should contain middles and limits of Cauchy sequences (identified with their defining Cauchy sequence). Thus, we consider a datatype containing points in the original space and closed under two operations: taking a pair of points in the datatype (we think of the resulting pair as the middle of the pair) and taking a sequence with values in the datatype (we think of the resulting sequence as the limit of the sequence if it is Cauchy, for a distance yet to be defined, and as something we discard if the sequence is not Cauchy).

Defining such an object is apparently not trivial. However, it is well defined, for cardinality reasons, as this process will end after the continuum cardinality iterations (as a sequence taking value in the continuum cardinality is in fact contained in a strictly smaller ordinal, which means that all sequences in the construction will appear at a step strictly before the continuum cardinality). The datatype construction in Isabelle/HOL contains these cardinality considerations as an automatic process, and is thus able to construct the datatype directly, without the need for any additional proof!

Then, we define a wellorder on the datatype, such that every middle and every sequence appear after each of its ancestors. This construction of a wellorder should work for any datatype, but we provide a naive proof in our use case.

Then, we define, inductively on z, a pseudodistance on the pair of points in $\{x : x \leq z\}$. In the induction, one should add one point at a time. If it is a middle, one uses the Bonk-Schramm recipe. If it is a sequence, then either the sequence is Cauchy and one uses the limit of the distances to the points in the sequence, or it is not Cauchy and one discards the new point by setting $d(a, a) = 1$. (This means that, in the Bonk-Schramm recipe, we only use the points with $d(x, x) = 0$, and show the triangular inequality there).

In the end, we obtain a space with a pseudodistance. The desired space is obtained by quotienting out the space $\{x : d(x, x) = 0\}$ by the equivalence relation given by $d(x, y) = 0$. The triangular inequality for the pseudodistance shows that it descends to a genuine distance on the quotient. This is the desired geodesic complete extension of the original space.

12.1 Unfolded Bonk Schramm extension

The unfolded Bonk Schramm extension, as explained at the beginning of this file, is a type made of the initial type, adding all possible middles and all possible limits of Cauchy sequences, without any quotienting process

datatype *'a Bonk-Schramm-extension-unfolded =*
  basepoint 'a
  | *middle 'a Bonk-Schramm-extension-unfolded 'a Bonk-Schramm-extension-unfolded*
  | *would-be-Cauchy nat ⇒ 'a Bonk-Schramm-extension-unfolded*

context *metric-space*
begin

The construction of the distance will be done by transfinite induction, with respect to a well-order for which the basepoints form an initial segment, and for which middles of would-be Cauchy sequences are larger than the elements they are made of. We will first prove the existence of such a well-order.

The idea is first to construct a function `map_aux` to another type, with a well-order `wo_aux`, such that the image of `middle a b` is larger than the images of `a` and `b` (take for instance the successor of the maximum of the two), and likewise for a Cauchy sequence. A definition by induction works if the target cardinal is large enough.

Then, pullback the well-order `wo_aux` by the map `map_aux`: this gives a relation that satisfies all the required properties, except that two different elements can be equal for the order. Extending it essentially arbitrarily to distinguish between all elements (this is done in Lemma `Well_order_pullback`) gives the desired well-order

definition *Bonk-Schramm-extension-unfolded-wo* **where**
Bonk-Schramm-extension-unfolded-wo = (*SOME* (*r*::'*a Bonk-Schramm-extension-unfolded rel*).
well-order-on UNIV r
∧ (∀ *x* ∈ *range basepoint*. ∀ *y* ∈ − *range basepoint*. (*x*, *y*) ∈ *r*)
∧ (∀ *a b*. (*a*, *middle a b*) ∈ *r*)
∧ (∀ *a b*. (*b*, *middle a b*) ∈ *r*)
∧ (∀ *u n*. (*u n*, *would-be-Cauchy u*) ∈ *r*))

We prove the existence of the well order

definition *wo-aux* **where**
wo-aux = (*SOME* (*r*:: (*nat* + '*a Bonk-Schramm-extension-unfolded set*) *rel*).
Card-order r ∧ ¬*finite*(*Field r*) ∧ *regularCard r* ∧ |*UNIV*::'*a Bonk-Schramm-extension-unfolded set*| <*o r*)

lemma *wo-aux-exists*:
Card-order wo-aux ∧ ¬*finite* (*Field wo-aux*) ∧ *regularCard wo-aux* ∧ |*UNIV*::'*a Bonk-Schramm-extension-unfolded set*| <*o wo-aux*
⟨*proof*⟩

interpretation *wo-aux*: *wo-rel wo-aux*
⟨*proof*⟩

primrec *map-aux*::'*a Bonk-Schramm-extension-unfolded* ⇒ *nat* + '*a Bonk-Schramm-extension-unfolded set* **where**
map-aux (*basepoint x*) = *wo-aux.zero*
| *map-aux* (*middle a b*) = *wo-aux.suc* ({*map-aux a*} ∪ {*map-aux b*})
| *map-aux* (*would-be-Cauchy u*) = *wo-aux.suc* ((*map-aux o u*)'*UNIV*)

lemma *map-aux-AboveS-not-empty*:

assumes *map-aux‘S* ⊆ *Field wo-aux*
shows *wo-aux.AboveS* (*map-aux‘S*) ≠ {}
⟨*proof*⟩

lemma *map-aux-in-Field*:
map-aux x ∈ *Field wo-aux*
⟨*proof*⟩

lemma *middle-rel-a*:
(*map-aux a*, *map-aux* (*middle a b*)) ∈ *wo-aux* − *Id*
⟨*proof*⟩

lemma *middle-rel-b*:
(*map-aux b*, *map-aux* (*middle a b*)) ∈ *wo-aux* − *Id*
⟨*proof*⟩

lemma *cauchy-rel*:
(*map-aux* (*u n*), *map-aux* (*would-be-Cauchy u*)) ∈ *wo-aux* − *Id*
⟨*proof*⟩

From the above properties of `wo_aux`, it follows using `Well_order_pullback` that an order satisfying all the properties we want of `Bonk_Schramm_extension_unfolded_wo` exists. Hence, we get the following lemma.

lemma *Bonk-Schramm-extension-unfolded-wo-props*:
well-order-on UNIV Bonk-Schramm-extension-unfolded-wo
∀ *x* ∈ *range basepoint*. ∀ *y* ∈ − *range basepoint*. (*x*, *y*) ∈ *Bonk-Schramm-extension-unfolded-wo*
∀ *a b*. (*a*, *middle a b*) ∈ *Bonk-Schramm-extension-unfolded-wo*
∀ *a b*. (*b*, *middle a b*) ∈ *Bonk-Schramm-extension-unfolded-wo*
∀ *u n*. (*u n*, *would-be-Cauchy u*) ∈ *Bonk-Schramm-extension-unfolded-wo*
⟨*proof*⟩

interpretation *wo*: *wo-rel Bonk-Schramm-extension-unfolded-wo*
⟨*proof*⟩

We reformulate in the interpretation `wo` the main properties of `Bonk_Schramm_extension_unfolded_wo` that we established in Lemma `Bonk_Schramm_extension_unfolded_wo_props`

lemma *Bonk-Schramm-extension-unfolded-wo-props'*:
a ∈ *wo.underS* (*middle a b*)
b ∈ *wo.underS* (*middle a b*)
u n ∈ *wo.underS* (*would-be-Cauchy u*)
⟨*proof*⟩

We want to define by transfinite induction a distance on `'a Bonk_Schramm_extension_unfolded`, adding one point at a time (i.e., if the distance is defined on E, then one wants to define it on $E \cup \{x\}$, if x is a middle or a potential Cauchy sequence, by prescribing the distance from x to all the points in E.

Technically, we define a family of distances, indexed by x, on $\{y : y \leq x\}^2$. As all functions should be defined everywhere, this will be a family of functions on $X \times X$, indexed by points in X. They will have a compatibility condition, making it possible to define a global distance by gluing them together.

Technically, transfinite induction is implemented in Isabelle/HOL by an updating rule: a function that associates, to a family of distances indexed by x, a new family of distances indexed by x. The result of the transfinite induction is obtained by starting from an arbitrary object, and then applying the updating rule infinitely many times. The characteristic property of the result of this transfinite induction is that it is a fixed point of the updating rule, as it should.

Below, this is implemented as follows:

- `extend_distance` is the updating rule.
- Its fixed point `extend_distance_fp` is by definition `wo.worec extend_distance` (it only makes sense if the udpating rule satisfies a compatibility condition `wo.adm_wo extend_distance` saying that the update of a family, at x, only depends on the value of the family strictly below x.
- Finally, the global distance `extended_distance` is taken as the value of the fixed point above, on xyy' (i.e., using the distance indexed by x) for any $x \geq \max(y, y')$. For definiteness, we use $\max(y, y')$, but it does not matter as everything is compatible.

fun *extend-distance*::(*'a Bonk-Schramm-extension-unfolded* ⇒ (*'a Bonk-Schramm-extension-unfolded* ⇒ *'a Bonk-Schramm-extension-unfolded* ⇒ *real*))
⇒ (*'a Bonk-Schramm-extension-unfolded* ⇒ (*'a Bonk-Schramm-extension-unfolded* ⇒ *'a Bonk-Schramm-extension-unfolded* ⇒ *real*))
where
extend-distance f (*basepoint x*) = (*λy z. if y* ∈ *range basepoint* ∧ *z* ∈ *range basepoint then*
dist (*SOME y'. y* = *basepoint y'*) (*SOME z'. z* = *basepoint z'*) *else 1*)
| *extend-distance f* (*middle a b*) = (*λy z.*
if (*y* ∈ *wo.underS* (*middle a b*)) ∧ (*z* ∈ *wo.underS* (*middle a b*)) *then f* (*wo.max2 y z*) *y z*
else if (*y* ∈ *wo.underS* (*middle a b*)) ∧ (*z* = *middle a b*) *then* (*f* (*wo.max2 a b*) *a b*)/*2* + (*SUP w*∈{*z* ∈ *wo.underS* (*middle a b*). *f z z z* = *0*}. *f* (*wo.max2 y w*) *y w* − *max* (*f* (*wo.max2 a w*) *a w*) (*f* (*wo.max2 b w*) *b w*))
else if (*y* = *middle a b*) ∧ (*z* ∈ *wo.underS* (*middle a b*)) *then* (*f* (*wo.max2 a b*) *a b*)/*2* + (*SUP w*∈{*z* ∈ *wo.underS* (*middle a b*). *f z z z* = *0*}. *f* (*wo.max2 z w*) *z w* − *max* (*f* (*wo.max2 a w*) *a w*) (*f* (*wo.max2 b w*) *b w*))
else if (*y* = *middle a b*) ∧ (*z* = *middle a b*) ∧ (*f a a a* = *0*) ∧ (*f b b b* = *0*) *then 0*
else 1)
| *extend-distance f* (*would-be-Cauchy u*) = (*λy z.*
if (*y* ∈ *wo.underS* (*would-be-Cauchy u*)) ∧ (*z* ∈ *wo.underS* (*would-be-Cauchy u*)) *then f* (*wo.max2 y z*) *y z*

else if $(\neg(\forall\, eps > (0::real).\ \exists\, N.\ \forall\, n \geq N.\ \forall\, m \geq N.\ f\ (wo.max2\ (u\ n)\ (u\ m))\ (u\ n)\ (u\ m) < eps))$ *then 1*
else if $(y \in wo.underS\ (would\text{-}be\text{-}Cauchy\ u)) \wedge (z = would\text{-}be\text{-}Cauchy\ u)$ *then* $lim\ (\lambda n.\ f\ (wo.max2\ (u\ n)\ y)\ (u\ n)\ y)$
else if $(y = would\text{-}be\text{-}Cauchy\ u) \wedge (z \in wo.underS\ (would\text{-}be\text{-}Cauchy\ u))$ *then* $lim\ (\lambda n.\ f\ (wo.max2\ (u\ n)\ z)\ (u\ n)\ z)$
else if $(y = would\text{-}be\text{-}Cauchy\ u) \wedge (z = would\text{-}be\text{-}Cauchy\ u) \wedge (\forall\, n.\ f\ (u\ n)\ (u\ n)\ (u\ n) = 0)$ *then 0*
else 1)

definition *extend-distance-fp* = *wo.worec extend-distance*

definition *extended-distance x y* = *extend-distance-fp* (*wo.max2 x y*) *x y*

definition *extended-distance-set* = {*z. extended-distance z z* = *0*}

lemma *wo-adm-extend-distance*:
wo.adm-wo extend-distance
⟨*proof*⟩

lemma *extend-distance-fp*:
extend-distance-fp = *extend-distance* (*extend-distance-fp*)
⟨*proof*⟩

lemma *extended-distance-symmetric*:
extended-distance x y = *extended-distance y x*
⟨*proof*⟩

lemma *extended-distance-basepoint*:
extended-distance (*basepoint x*) (*basepoint y*) = *dist x y*
⟨*proof*⟩

lemma *extended-distance-set-basepoint*:
basepoint x ∈ *extended-distance-set*
⟨*proof*⟩

lemma *extended-distance-set-middle*:
assumes *a* ∈ *extended-distance-set b* ∈ *extended-distance-set*
shows *middle a b* ∈ *extended-distance-set*
⟨*proof*⟩

lemma *extended-distance-set-middle'*:
assumes *middle a b* ∈ *extended-distance-set*
shows *a* ∈ *extended-distance-set* ∩ *wo.underS* (*middle a b*)
b ∈ *extended-distance-set* ∩ *wo.underS* (*middle a b*)
⟨*proof*⟩

lemma *extended-distance-middle-formula*:
assumes *x* ∈ *wo.underS* (*middle a b*)

shows *extended-distance x* (*middle a b*) = (*extended-distance a b*)/*2*
+ (*SUP w*∈*wo.underS* (*middle a b*) ∩ *extended-distance-set*.
extended-distance x w − *max* (*extended-distance a w*) (*extended-distance b w*))
⟨*proof*⟩

lemma *extended-distance-set-Cauchy*:
assumes (*would-be-Cauchy u*) ∈ *extended-distance-set*
shows *u n* ∈ *extended-distance-set* ∩ *wo.underS* (*would-be-Cauchy u*)
∀ *eps* > (*0*::*real*). ∃ *N*. ∀ *n* ≥ *N*. ∀ *m* ≥ *N*. *extended-distance* (*u n*) (*u m*) < *eps*
⟨*proof*⟩

lemma *extended-distance-triang-ineq*:
assumes *x* ∈ *extended-distance-set*
y ∈ *extended-distance-set*
z ∈ *extended-distance-set*
shows *extended-distance x z* ≤ *extended-distance x y* + *extended-distance y z*
⟨*proof*⟩

We can now show the two main properties of the construction: the middle is indeed a middle from the metric point of view (in `extended_distance_middle`), and Cauchy sequences have a limit (the corresponding `would_be_Cauchy` point).

lemma *extended-distance-pos*:
assumes *a* ∈ *extended-distance-set*
b ∈ *extended-distance-set*
shows *extended-distance a b* ≥ *0*
⟨*proof*⟩

lemma *extended-distance-middle*:
assumes *a* ∈ *extended-distance-set*
b ∈ *extended-distance-set*
shows *extended-distance a* (*middle a b*) = *extended-distance a b* / *2*
extended-distance b (*middle a b*) = *extended-distance a b* / *2*
⟨*proof*⟩

lemma *extended-distance-Cauchy*:
assumes ⋀(*n*::*nat*). *u n* ∈ *extended-distance-set*
and ∀ *eps* > (*0*::*real*). ∃ *N*. ∀ *n* ≥ *N*. ∀ *m* ≥ *N*. *extended-distance* (*u n*) (*u m*) < *eps*
shows *would-be-Cauchy u* ∈ *extended-distance-set*
(λ*n*. *extended-distance* (*u n*) (*would-be-Cauchy u*)) ⟶ *0*
⟨*proof*⟩

end

12.2 The Bonk Schramm extension

quotient-type (**overloaded**) *'a Bonk-Schramm-extension* =
(*'a*::*metric-space*) *Bonk-Schramm-extension-unfolded*
/ *partial*: $\lambda x\ y.\ (x \in$ *extended-distance-set* $\wedge\ y \in$ *extended-distance-set* $\wedge$ *extended-distance* $x\ y = 0)$
⟨*proof*⟩

instantiation *Bonk-Schramm-extension* :: (*metric-space*) *metric-space*
begin

lift-definition *dist-Bonk-Schramm-extension*::(*'a*::*metric-space*) *Bonk-Schramm-extension* ⇒ *'a Bonk-Schramm-extension* ⇒ *real*
is $\lambda x\ y.$ *extended-distance* $x\ y$
⟨*proof*⟩

To define a metric space in the current library of Isabelle/HOL, one should also introduce a uniformity structure and a topology, as follows (they are prescribed by the distance):

definition *uniformity-Bonk-Schramm-extension*::((*'a Bonk-Schramm-extension*) × (*'a Bonk-Schramm-extension*)) *filter*
where *uniformity-Bonk-Schramm-extension* = (*INF* $e \in \{0 <..\}$. *principal* $\{(x, y).$ *dist* $x\ y < e\})$

definition *open-Bonk-Schramm-extension* :: *'a Bonk-Schramm-extension set* ⇒ *bool*
where *open-Bonk-Schramm-extension* $U = (\forall x \in U.$ *eventually* $(\lambda(x', y).\ x' = x \longrightarrow y \in U)$ *uniformity*)

instance ⟨*proof*⟩
end

instance *Bonk-Schramm-extension* :: (*metric-space*) *complete-space*
⟨*proof*⟩

instance *Bonk-Schramm-extension* :: (*metric-space*) *geodesic-space*
⟨*proof*⟩

definition *to-Bonk-Schramm-extension*::*'a*::*metric-space* ⇒ *'a Bonk-Schramm-extension*
where *to-Bonk-Schramm-extension* x = *abs-Bonk-Schramm-extension* (*basepoint* x)

lemma *to-Bonk-Schramm-extension-isometry*:
isometry-on UNIV to-Bonk-Schramm-extension
⟨*proof*⟩

13 Bonk-Schramm extension of hyperbolic spaces

13.1 The Bonk-Schramm extension preserves hyperbolicity

A central feature of the Bonk-Schramm extension is that it preserves hyperbolicity, with the same hyperbolicity constant δ, as we prove now.

lemma (**in** *Gromov-hyperbolic-space*) *Bonk-Schramm-extension-unfolded-hyperbolic*:
fixes *x y z t*::(*'a*::*metric-space*) *Bonk-Schramm-extension-unfolded*
assumes *x* ∈ *extended-distance-set*
y ∈ *extended-distance-set*
z ∈ *extended-distance-set*
t ∈ *extended-distance-set*
shows *extended-distance x y* + *extended-distance z t* ≤ *max* (*extended-distance x z* + *extended-distance y t*) (*extended-distance x t* + *extended-distance y z*) + *2* ∗ *deltaG*(*TYPE*(*'a*))
⟨*proof*⟩

lemma (**in** *Gromov-hyperbolic-space*) *Bonk-Schramm-extension-hyperbolic*:
Gromov-hyperbolic-subset (*deltaG*(*TYPE*(*'a*))) (*UNIV*::(*'a Bonk-Schramm-extension*) *set*)
⟨*proof*⟩

instantiation *Bonk-Schramm-extension* :: (*Gromov-hyperbolic-space*) *Gromov-hyperbolic-space-geodesic*
begin
definition *deltaG-Bonk-Schramm-extension*::(*'a Bonk-Schramm-extension*) *itself* ⇒ *real* **where**
deltaG-Bonk-Schramm-extension - = *deltaG*(*TYPE*(*'a*))

instance ⟨*proof*⟩
end

Finally, it follows that the Bonk Schramm extension of a 0-hyperbolic space (in which it embeds isometrically) is a metric tree or, equivalently, a geodesic 0-hyperbolic space (the equivalence is proved at the end of `Geodesic_Spaces.thy`).

instance *Bonk-Schramm-extension* :: (*Gromov-hyperbolic-space-0*) *Gromov-hyperbolic-space-0-geodesic*
⟨*proof*⟩

It then follows that it is also a metric tree, from what we have already proved. We write explicitly for definiteness.

instance *Bonk-Schramm-extension* :: (*Gromov-hyperbolic-space-0*) *metric-tree*
⟨*proof*⟩

13.2 Applications

We deduce that we can extend results on Gromov-hyperbolic spaces without the geodesicity assumption, even if it is used in the proofs. These results are given for illustrative purpose mainly, as one works most often in geodesic spaces anyway.

The following results have already been proved in hyperbolic geodesic spaces. The same results follow in a general hyperbolic space, as everything is invariant under isometries and can thus be pulled from the corresponding result in the Bonk Schramm extension. The straightforward proofs only express this invariance under isometries of all the properties under consideration.

proposition (**in** *Gromov-hyperbolic-space*) *lipschitz-path-close-to-geodesic'*:
  fixes *c*::*real* ⇒ *'a*
  assumes *lipschitz-on M {A..B} c*
    geodesic-segment-between G (c A) (c B)
    x ∈ *G*
  shows *infdist x (c'{A..B})* ≤ *(4/ln 2) ∗ deltaG(TYPE('a)) ∗ max 0 (ln (B−A))* + *M*
⟨*proof*⟩

theorem (**in** *Gromov-hyperbolic-space*) *Morse-Gromov-theorem'*:
  fixes *f*::*real* ⇒ *'a*
  assumes *lambda C−quasi-isometry-on {a..b} f*
    geodesic-segment-between G (f a) (f b)
  shows *hausdorff-distance (f'{a..b}) G* ≤ *92 ∗ lambda*^{2} *∗ (C + deltaG(TYPE('a)))*
⟨*proof*⟩

theorem (**in** *Gromov-hyperbolic-space*) *Morse-Gromov-theorem2'*:
  fixes *c d*::*real* ⇒ *'a*
  assumes *lambda C−quasi-isometry-on {A..B} c*
    lambda C−quasi-isometry-on {A..B} d
    c A = d A c B = d B
  shows *hausdorff-distance (c'{A..B}) (d'{A..B})* ≤ *184 ∗ lambda^2 ∗ (C + deltaG(TYPE('a)))*
⟨*proof*⟩

lemma *Gromov-hyperbolic-invariant-under-quasi-isometry-explicit'*:
  fixes *f*::*'a*::*geodesic-space* ⇒ *'b*::*Gromov-hyperbolic-space*
  assumes *lambda C−quasi-isometry f*
  shows *Gromov-hyperbolic-subset (752 ∗ lambda^3 ∗ (C + deltaG(TYPE('b)))) (UNIV::('a set))*
⟨*proof*⟩

theorem *Gromov-hyperbolic-invariant-under-quasi-isometry'*:
  assumes *quasi-isometric (UNIV::('a::geodesic-space) set) (UNIV::('b::Gromov-hyperbolic-space) set)*
  shows ∃ *delta. Gromov-hyperbolic-subset delta (UNIV::'a set)*
⟨*proof*⟩

end

theory *Gromov-Boundary*
  imports *Gromov-Hyperbolicity Eexp-Eln*

begin

14 Constructing a distance from a quasi-distance

Below, we will construct a distance on the Gromov completion of a hyperbolic space. The geometrical object that arises naturally is almost a distance, but it does not satisfy the triangular inequality. There is a general process to turn such a quasi-distance into a genuine distance, as follows: define the new distance $\tilde{d}(x, y)$ to be the infimum of $d(x, u_1) + d(u_1, u_2) + \cdots + d(u_{n-1}, x)$ over all sequences of points (of any length) connecting x to y. It is clear that it satisfies the triangular inequality, is symmetric, and $\tilde{d}(x, y) \leq d(x, y)$. What is not clear, however, is if $\tilde{d}(x, y)$ can be zero if $x \neq y$, or more generally how one can bound $\tilde{d}$ from below. The main point of this contruction is that, if d satisfies the inequality $d(x, z) \leq \sqrt{2} \max(d(x, y), d(y, z))$, then one has $\tilde{d}(x, y) \geq d(x, y)/2$ (and in particular $\tilde{d}$ defines the same topology, the same set of Lipschitz functions, and so on, as d).

This statement can be found in [Bourbaki, topologie generale, chapitre 10] or in [Ghys-de la Harpe] for instance. We follow their proof.

definition *turn-into-distance*::(*'a* ⇒ *'a* ⇒ *real*) ⇒ (*'a* ⇒ *'a* ⇒ *real*)
  where *turn-into-distance f x y* = *Inf* {(∑ *i* ∈ {*0*..<*n*}. *f* (*u i*) (*u* (*Suc i*)))| *u* (*n*::*nat*). *u 0* = *x* ∧ *u n* = *y*}

locale *Turn-into-distance* =
  fixes *f*::*'a* ⇒ *'a* ⇒ *real*
  assumes *nonneg*: *f x y* ≥ *0*
    and *sym*: *f x y* = *f y x*
    and *self-zero*: *f x x* = *0*
    and *weak-triangle*: *f x z* ≤ *sqrt 2* ∗ *max* (*f x y*) (*f y z*)
begin

The two lemmas below are useful when dealing with Inf results, as they always require the set under consideration to be non-empty and bounded from below.

lemma *bdd-below* [*simp*]:
  bdd-below {(∑ *i* = *0*..<*n*. *f* (*u i*) (*u* (*Suc i*)))| *u* (*n*::*nat*). *u 0* = *x* ∧ *u n* = *y*}
  ⟨*proof*⟩

lemma *nonempty*:
  {∑ *i* = *0*..<*n*. *f* (*u i*) (*u* (*Suc i*)) |*u n*. *u 0* = *x* ∧ *u n* = *y*} ≠ {}
⟨*proof*⟩

We can now prove that `turn_into_distance f` satisfies all the properties of a distance. First, it is nonnegative.

lemma *TID-nonneg*:
  turn-into-distance f x y ≥ *0*

⟨*proof*⟩

For the symmetry, we use the symmetry of f, and go backwards along a chain of points, replacing a sequence from x to y with a sequence from y to x.

lemma *TID-sym*:
turn-into-distance f x y = turn-into-distance f y x
⟨*proof*⟩

There is a trivial upper bound by f, using the single chain x, y.

lemma *upper*:
turn-into-distance f x y ≤ f x y
⟨*proof*⟩

The new distance vanishes on a pair of equal points, as this is already the case for f.

lemma *TID-self-zero*:
turn-into-distance f x x = 0
⟨*proof*⟩

For the triangular inequality, we concatenate a sequence from x to y almost realizing the infimum, and a sequence from y to z almost realizing the infimum, to obtain a sequence from x to z along which the sums of f is almost bounded by `turn_into_distance f x y + turn_into_distance f y z`.

lemma *triangle*:
turn-into-distance f x z ≤ turn-into-distance f x y + turn-into-distance f y z
⟨*proof*⟩

Now comes the only nontrivial statement of the construction, the fact that the new distance is bounded from below by $f/2$.

Here is the mathematical proof. We show by induction that all chains from x to y satisfy this bound. Assume this is done for all chains of length $< n$, we do it for a chain of length n. Write $S = \sum f(u_i, u_{i+1})$ for the sum along the chain. Introduce p the last index where the sum is $\leq S/2$. Then the sum from 0 to p is $\leq S/2$, and the sum from $p+1$ to n is also $\leq S/2$ (by maximality of p). The induction assumption gives that $f(x, u_p)$ is bounded by twice the sum from 0 to p, which is at most S. Same thing for $f(u_{p+1}, y)$. With the weird triangle inequality applied two times, we get $f(x, y) \leq 2 \max(f(x, u_p), f(u_p, u_{p+1}), f(u_{p+1}, y)) \leq 2S$, as claimed.

The formalization presents no difficulty.

lemma *lower*:
f x y ≤ 2 ∗ turn-into-distance f x y
⟨*proof*⟩

end

15 The Gromov completion of a hyperbolic space

15.1 The Gromov boundary as a set

A sequence in a Gromov hyperbolic space converges to a point in the boundary if the Gromov product $(u_n, u_m)_e$ tends to infinity when $m, n \to_i nfty$. The point at infinity is defined as the equivalence class of such sequences, for the relation $u \sim v$ iff $(u_n, v_n)_e \to \infty$ (or, equivalently, $(u_n, v_m)_e \to \infty$ when $m, n \to \infty$, or one could also change basepoints). Hence, the Gromov boundary is naturally defined as a quotient type. There is a difficulty: it can be empty in general, hence defining it as a type is not always possible. One could introduce a new typeclass of Gromov hyperbolic spaces for which the boundary is not empty (unboundedness is not enough, think of infinitely many segments $[0, n]$ all joined at 0), and then only define the boundary of such spaces. However, this is tedious. Rather, we work with the Gromov completion (containing the space and its boundary), this is always not empty. The price to pay is that, in the definition of the completion, we have to distinguish between sequences converging to the boundary and sequences converging inside the space. This is more natural to proceed in this way as the interesting features of the boundary come from the fact that its sits at infinity of the initial space, so their relations (and the topology of $X \cup \partial X$) are central.

definition *Gromov-converging-at-boundary*::(*nat* ⇒ (*'a*::*Gromov-hyperbolic-space*)) ⇒ *bool*
  where *Gromov-converging-at-boundary u* = (∀ *a*. ∀ (*M*::*real*). ∃ *N*. ∀ *n* ≥ *N*. ∀ *m* ≥ *N*. *Gromov-product-at a* (*u m*) (*u n*) ≥ *M*)

lemma *Gromov-converging-at-boundaryI*:
  assumes ⋀*M*. ∃ *N*. ∀ *n* ≥ *N*. ∀ *m* ≥ *N*. *Gromov-product-at a* (*u m*) (*u n*) ≥ *M*
  shows *Gromov-converging-at-boundary u*
⟨*proof*⟩

lemma *Gromov-converging-at-boundary-imp-unbounded*:
  assumes *Gromov-converging-at-boundary u*
  shows (λ*n*. *dist a* (*u n*)) ⟶ ∞
⟨*proof*⟩

lemma *Gromov-converging-at-boundary-imp-not-constant*:
  ¬(*Gromov-converging-at-boundary* (λ*n*. *x*))
  ⟨*proof*⟩

lemma *Gromov-converging-at-boundary-imp-not-constant'*:
  assumes *Gromov-converging-at-boundary u*
  shows ¬(∀ *m n*. *u m* = *u n*)
  ⟨*proof*⟩

We introduce a partial equivalence relation, defined over the sequences that converge to infinity, and the constant sequences. Quotienting the space of admissible sequences by this equivalence relation will give rise to the Gromov completion.

definition *Gromov-completion-rel*::(*nat* ⇒ *'a*::*Gromov-hyperbolic-space*) ⇒ (*nat* ⇒ *'a*) ⇒ *bool*
where *Gromov-completion-rel u v* =
(((*Gromov-converging-at-boundary u* ∧ *Gromov-converging-at-boundary v* ∧ (∀ *a*. (λ*n*. *Gromov-product-at a* (*u n*) (*v n*)) ⟶ ∞)))
∨ (∀ *n m*. *u n* = *v m* ∧ *u n* = *u m* ∧ *v n* = *v m*))

We need some basic lemmas to work separately with sequences tending to the boundary and with constant sequences, as follows.

lemma *Gromov-completion-rel-const* [*simp*]:
Gromov-completion-rel (λ*n*. *x*) (λ*n*. *x*)
⟨*proof*⟩

lemma *Gromov-completion-rel-to-const*:
assumes *Gromov-completion-rel u* (λ*n*. *x*)
shows *u n* = *x*
⟨*proof*⟩

lemma *Gromov-completion-rel-to-const'*:
assumes *Gromov-completion-rel* (λ*n*. *x*) *u*
shows *u n* = *x*
⟨*proof*⟩

lemma *Gromov-product-tendsto-PInf-a-b*:
assumes (λ*n*. *Gromov-product-at a* (*u n*) (*v n*)) ⟶ ∞
shows (λ*n*. *Gromov-product-at b* (*u n*) (*v n*)) ⟶ ∞
⟨*proof*⟩

lemma *Gromov-converging-at-boundary-rel*:
assumes *Gromov-converging-at-boundary u*
shows *Gromov-completion-rel u u*
⟨*proof*⟩

We can now prove that we indeed have an equivalence relation.

lemma *part-equivp-Gromov-completion-rel*:
part-equivp Gromov-completion-rel
⟨*proof*⟩

We can now define the Gromov completion of a Gromov hyperbolic space, considering either sequences converging to a point on the boundary, or sequences converging inside the space, and quotienting by the natural equivalence relation.

quotient-type (**overloaded**) *'a Gromov-completion* =
nat ⇒ (*'a*::*Gromov-hyperbolic-space*)

/ *partial*: *Gromov-completion-rel*
⟨*proof*⟩

The Gromov completion contains is made of a copy of the original space, and new points forming the Gromov boundary.

definition *to-Gromov-completion*::(*'a*::*Gromov-hyperbolic-space*) ⇒ *'a Gromov-completion*
  where *to-Gromov-completion x* = *abs-Gromov-completion* (*λn. x*)

definition *from-Gromov-completion*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion* ⇒ *'a*
  where *from-Gromov-completion* = *inv to-Gromov-completion*

definition *Gromov-boundary*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion set*
  where *Gromov-boundary* = *UNIV* − *range to-Gromov-completion*

lemma *to-Gromov-completion-inj*:
  inj to-Gromov-completion
⟨*proof*⟩

lemma *from-to-Gromov-completion* [*simp*]:
  from-Gromov-completion (*to-Gromov-completion x*) = *x*
⟨*proof*⟩

lemma *to-from-Gromov-completion*:
  assumes *x* ∉ *Gromov-boundary*
  shows *to-Gromov-completion* (*from-Gromov-completion x*) = *x*
⟨*proof*⟩

lemma *not-in-Gromov-boundary*:
  assumes *x* ∉ *Gromov-boundary*
  shows ∃ *a*. *x* = *to-Gromov-completion a*
⟨*proof*⟩

lemma *not-in-Gromov-boundary'* [*simp*]:
  to-Gromov-completion x ∉ *Gromov-boundary*
⟨*proof*⟩

lemma *abs-Gromov-completion-in-Gromov-boundary* [*simp*]:
  assumes *Gromov-converging-at-boundary u*
  shows *abs-Gromov-completion u* ∈ *Gromov-boundary*
⟨*proof*⟩

lemma *rep-Gromov-completion-to-Gromov-completion* [*simp*]:
  rep-Gromov-completion (*to-Gromov-completion y*) = (*λn. y*)
⟨*proof*⟩

To distinguish the case of points inside the space or in the boundary, we introduce the following case distinction.

lemma *Gromov-completion-cases* [*case-names to-Gromov-completion boundary*, *cases*

type: *Gromov-completion*]:
  ($\bigwedge x$. z = *to-Gromov-completion* $x \Longrightarrow P$) $\Longrightarrow$ ($z \in$ *Gromov-boundary* $\Longrightarrow P$) $\Longrightarrow P$
⟨*proof*⟩

15.2 Extending the original distance and the original Gromov product to the completion

In this subsection, we extend the Gromov product to the boundary, by taking limits along sequences tending to the point in the boundary. This does not converge, but it does up to δ, so for definiteness we use a $\liminf$ over all sequences tending to the boundary point – one interest of this definition is that the extended Gromov product still satisfies the hyperbolicity inequality. One difficulty is that this extended Gromov product can take infinite values (it does so exactly on the pair (x, x) where x is in the boundary), so we should define this product in extended nonnegative reals.

We also extend the original distance, by $+\infty$ on the boundary. This is not a really interesting function, but it will be instrumental below. Again, this extended Gromov distance (not to be mistaken for the genuine distance we will construct later on on the completion) takes values in extended nonnegative reals.

Since the extended Gromov product and the extension of the original distance both take values in $[0, +\infty]$, it may seem natural to define them in ennreal. This is the choice that was made in a previous implementation, but it turns out that one keeps computing with these numbers, writing down inequalities and subtractions. ennreal is ill suited for this kind of computations, as it only works well with additions. Hence, the implementation was switched to ereal, where proofs are indeed much smoother.

To define the extended Gromov product, one takes a limit of the Gromov product along any sequence, as it does not depend up to δ on the chosen sequence. However, if one wants to keep the exact inequality that defines hyperbolicity, but at all points, then using an infimum is the best choice.

definition *extended-Gromov-product-at*::(*'a*::*Gromov-hyperbolic-space*) ⇒ *'a Gromov-completion* ⇒ *'a Gromov-completion* ⇒ *ereal*
 where *extended-Gromov-product-at e x y* = *Inf* {*liminf* (*λn*. *ereal*(*Gromov-product-at e* (*u n*) (*v n*))) |*u v*. *abs-Gromov-completion u* = *x* ∧ *abs-Gromov-completion v* = *y* ∧ *Gromov-completion-rel u u* ∧ *Gromov-completion-rel v v*}

definition *extended-Gromov-distance*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion* ⇒ *'a Gromov-completion* ⇒ *ereal*
  where *extended-Gromov-distance x y* =
    (*if x* ∈ *Gromov-boundary* ∨ *y* ∈ *Gromov-boundary then* ∞
    else ereal (*dist* (*inv to-Gromov-completion x*) (*inv to-Gromov-completion y*)))

The extended distance and the extended Gromov product are invariant under exchange of the points, readily from the definition.

lemma *extended-Gromov-distance-commute*:
extended-Gromov-distance x y = extended-Gromov-distance y x
⟨*proof*⟩

lemma *extended-Gromov-product-nonneg* [*mono-intros*, *simp*]:
0 ≤ extended-Gromov-product-at e x y
⟨*proof*⟩

lemma *extended-Gromov-distance-nonneg* [*mono-intros*, *simp*]:
0 ≤ extended-Gromov-distance x y
⟨*proof*⟩

lemma *extended-Gromov-product-at-commute*:
extended-Gromov-product-at e x y = extended-Gromov-product-at e y x
⟨*proof*⟩

Inside the space, the extended distance and the extended Gromov product coincide with the original ones.

lemma *extended-Gromov-distance-inside* [*simp*]:
extended-Gromov-distance (*to-Gromov-completion x*) (*to-Gromov-completion y*) = *dist x y*
⟨*proof*⟩

lemma *extended-Gromov-product-inside* [*simp*] :
extended-Gromov-product-at e (*to-Gromov-completion x*) (*to-Gromov-completion y*) = *Gromov-product-at e x y*
⟨*proof*⟩

A point in the boundary is at infinite extended distance of everyone, including itself: the extended distance is obtained by taking the supremum along all sequences tending to this point, so even for one single point one can take two sequences tending to it at different speeds, which results in an infinite extended distance.

lemma *extended-Gromov-distance-PInf-boundary* [*simp*]:
assumes *x ∈ Gromov-boundary*
shows *extended-Gromov-distance x y = ∞ extended-Gromov-distance y x = ∞*
⟨*proof*⟩

By construction, the extended distance still satisfies the triangle inequality.

lemma *extended-Gromov-distance-triangle* [*mono-intros*]:
extended-Gromov-distance x z ≤ extended-Gromov-distance x y + extended-Gromov-distance y z
⟨*proof*⟩

The extended Gromov product can be bounded by the extended distance, just like inside the space.

lemma *extended-Gromov-product-le-dist* [*mono-intros*]:
extended-Gromov-product-at e x y $\leq$ *extended-Gromov-distance* (*to-Gromov-completion e*) *x*
⟨*proof*⟩

lemma *extended-Gromov-product-le-dist′* [*mono-intros*]:
extended-Gromov-product-at e x y $\leq$ *extended-Gromov-distance* (*to-Gromov-completion e*) *y*
⟨*proof*⟩

The Gromov product inside the space varies by at most the distance when one varies one of the points. We will need the same statement for the extended Gromov product. The proof is done using this inequality inside the space, and passing to the limit.

lemma *extended-Gromov-product-at-diff3* [*mono-intros*]:
extended-Gromov-product-at e x y $\leq$ *extended-Gromov-product-at e x z* + *extended-Gromov-distance y z*
⟨*proof*⟩

lemma *extended-Gromov-product-at-diff2* [*mono-intros*]:
extended-Gromov-product-at e x y $\leq$ *extended-Gromov-product-at e z y* + *extended-Gromov-distance x z*
⟨*proof*⟩

lemma *extended-Gromov-product-at-diff1* [*mono-intros*]:
extended-Gromov-product-at e x y $\leq$ *extended-Gromov-product-at f x y* + *dist e f*
⟨*proof*⟩

A point in the Gromov boundary is represented by a sequence tending to infinity and converging in the Gromov boundary, essentially by definition.

lemma *Gromov-boundary-abs-converging*:
assumes $x \in$ *Gromov-boundary abs-Gromov-completion u* = *x Gromov-completion-rel u u*
shows *Gromov-converging-at-boundary u*
⟨*proof*⟩

lemma *Gromov-boundary-rep-converging*:
assumes $x \in$ *Gromov-boundary*
shows *Gromov-converging-at-boundary* (*rep-Gromov-completion x*)
⟨*proof*⟩

We can characterize the points for which the Gromov product is infinite: they have to be the same point, at infinity. This is essentially equivalent to the definition of the Gromov completion, but there is some boilerplate to get the proof working.

lemma *Gromov-boundary-extended-product-PInf* [*simp*]:
extended-Gromov-product-at e x y = $\infty \longleftrightarrow$ ($x \in$ *Gromov-boundary* $\wedge$ *y* = *x*)
⟨*proof*⟩

As for points inside the space, we deduce that the extended Gromov product between x and x is just the extended distance to the basepoint.

lemma *extended-Gromov-product-e-x-x* [*simp*]:
extended-Gromov-product-at e x x = *extended-Gromov-distance* (*to-Gromov-completion e*) *x*
⟨*proof*⟩

The inequality in terms of Gromov products characterizing hyperbolicity extends in the same form to the Gromov completion, by taking limits of this inequality in the space.

lemma *extended-hyperb-ineq* [*mono-intros*]:
extended-Gromov-product-at (*e*::′*a*::*Gromov-hyperbolic-space*) *x z* ≥
min (*extended-Gromov-product-at e x y*) (*extended-Gromov-product-at e y z*) − *deltaG*(*TYPE*(′*a*))
⟨*proof*⟩

lemma *extended-hyperb-ineq′* [*mono-intros*]:
extended-Gromov-product-at (*e*::′*a*::*Gromov-hyperbolic-space*) *x z* + *deltaG*(*TYPE*(′*a*)) ≥
min (*extended-Gromov-product-at e x y*) (*extended-Gromov-product-at e y z*)
⟨*proof*⟩

lemma *zero-le-ereal* [*mono-intros*]:
assumes *0* ≤ *z*
shows *0* ≤ *ereal z*
⟨*proof*⟩

lemma *extended-hyperb-ineq-4-points′* [*mono-intros*]:
Min {*extended-Gromov-product-at* (*e*::′*a*::*Gromov-hyperbolic-space*) *x y*, *extended-Gromov-product-at e y z*, *extended-Gromov-product-at e z t*} ≤ *extended-Gromov-product-at e x t* + *2* ∗ *deltaG*(*TYPE*(′*a*))
⟨*proof*⟩

lemma *extended-hyperb-ineq-4-points* [*mono-intros*]:
Min {*extended-Gromov-product-at* (*e*::′*a*::*Gromov-hyperbolic-space*) *x y*, *extended-Gromov-product-at e y z*, *extended-Gromov-product-at e z t*} − *2* ∗ *deltaG*(*TYPE*(′*a*)) ≤ *extended-Gromov-product-at e x t*
⟨*proof*⟩

15.3 Construction of the distance on the Gromov completion

We want now to define the natural topology of the Gromov completion. Most textbooks first define a topology on ∂X, or sometimes on $X \cup \partial X$, and then much later a distance on ∂X (but they never do the tedious verification that the distance defines the same topology as the topology defined before). I have not seen one textbook defining a distance on $X \cup \partial X$. It turns out that one can in fact define a distance on $X \cup \partial X$, whose restriction to ∂X

is the usual distance on the Gromov boundary, and define the topology of $X \cup \partial X$ using it. For formalization purposes, this is very convenient as topologies defined with distances are automatically nice and tractable (no need to check separation axioms, for instance). The price to pay is that, once we have defined the distance, we have to check that it defines the right notion of convergence one expects.

What we would like to take for the distance is $d(x, y) = e^{-(x,y)_o}$, where o is some fixed basepoint in the space. However, this does not behave like a distance at small scales (but it is essentially the right thing at large scales), and it does not really satisfy the triangle inequality. However, $e^{-\epsilon(x,y)_o}$ almost satisfies the triangle inequality if ϵ is small enough, i.e., it is equivalent to a function satisfying the triangle inequality. This gives a genuine distance on the boundary, but not inside the space as it does not vanish on pairs (x, x).

A third try would be to take $d(x, y) = \min(\tilde{d}(x, y), e^{-\epsilon(x,y)_o})$ where $\tilde{d}$ is the natural extension of d to the Gromov completion (it is infinite if x or y belongs to the boundary). However, we can not prove that it is equivalent to a distance.

Finally, it works with $d(x, y) \asymp \min(\tilde{d}(x, y)^{1/2}, e^{-\epsilon(x,y)_o}$. This is what we will prove below. To construct the distance, we use the results proved in the locale `Turn_into_distance`. For this, we need to check that our quasi-distance satisfies a weird version of the triangular inequality.

All this construction depends on a basepoint, that we fix arbitrarily once and for all.

definition *epsilonG*::(*'a*::*Gromov-hyperbolic-space*) *itself* ⇒ *real*
  where *epsilonG - = ln 2 / (2+2∗deltaG(TYPE('a)))*

definition *basepoint*::*'a*
  where *basepoint = (SOME a. True)*

lemma *constant-in-extended-predist-pos* [*simp*, *mono-intros*]:
  epsilonG(TYPE('a::Gromov-hyperbolic-space)) > 0
  epsilonG(TYPE('a::Gromov-hyperbolic-space)) ≥ 0
  ennreal (epsilonG(TYPE('a))) ∗ top = top
⟨*proof*⟩

definition *extended-predist*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion* ⇒ *'a Gromov-completion* ⇒ *real*
 where *extended-predist x y = real-of-ereal (min (esqrt (extended-Gromov-distance x y))*
    (eexp (− epsilonG(TYPE('a)) ∗ extended-Gromov-product-at basepoint x y)))

lemma *extended-predist-ereal*:
  ereal (extended-predist x (y::('a::Gromov-hyperbolic-space) Gromov-completion)) = min (esqrt (extended-Gromov-distance x y))

(*eexp* (− *epsilonG*(*TYPE*(′*a*)) ∗ *extended-Gromov-product-at basepoint x y*))
⟨*proof*⟩

lemma *extended-predist-nonneg* [*simp*, *mono-intros*]:
extended-predist x y ≥ *0*
⟨*proof*⟩

lemma *extended-predist-commute*:
extended-predist x y = *extended-predist y x*
⟨*proof*⟩

lemma *extended-predist-self0* [*simp*]:
extended-predist x y = *0* ⟷ *x* = *y*
⟨*proof*⟩

lemma *extended-predist-le1* [*simp*, *mono-intros*]:
extended-predist x y ≤ *1*
⟨*proof*⟩

lemma *extended-predist-weak-triangle*:
extended-predist x z ≤ *sqrt 2* ∗ *max* (*extended-predist x y*) (*extended-predist y z*)
⟨*proof*⟩

instantiation *Gromov-completion* :: (*Gromov-hyperbolic-space*) *metric-space*
begin

definition *dist-Gromov-completion*::(′*a*::*Gromov-hyperbolic-space*) *Gromov-completion* ⇒ ′*a Gromov-completion* ⇒ *real*
where *dist-Gromov-completion* = *turn-into-distance extended-predist*

To define a metric space in the current library of Isabelle/HOL, one should also introduce a uniformity structure and a topology, as follows (they are prescribed by the distance):

definition *uniformity-Gromov-completion*::((′*a Gromov-completion*) × (′*a Gromov-completion*)) *filter*
where *uniformity-Gromov-completion* = (*INF e*∈{*0* <..}. *principal* {(*x*, *y*). *dist x y* < *e*})

definition *open-Gromov-completion* :: ′*a Gromov-completion set* ⇒ *bool*
where *open-Gromov-completion U* = (∀ *x*∈*U*. *eventually* (λ(*x*′, *y*). *x*′ = *x* ⟶ *y* ∈ *U*) *uniformity*)

instance ⟨*proof*⟩
end

The only relevant property of the distance on the Gromov completion is that it is comparable to the minimum of (the square root of) the extended distance, and the exponential of minus the Gromov product. The precise formula we use to define it is just an implementation detail, in a sense. We summarize these properties in the next theorem. From this point on, we will only use this, and never come back to the definition based on `extended_predist` and `turn_into_distance`.

theorem *Gromov-completion-dist-comparison* [*mono-intros*]:
  fixes *x y*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion*
  shows *ereal*(*dist x y*) $\leq$ *esqrt*(*extended-Gromov-distance x y*)
    ereal(*dist x y*) $\leq$ *eexp* (− *epsilonG*(*TYPE*(*'a*)) $*$ *extended-Gromov-product-at basepoint x y*)
    min (*esqrt*(*extended-Gromov-distance x y*)) (*eexp* (− *epsilonG*(*TYPE*(*'a*)) $*$ *extended-Gromov-product-at basepoint x y*)) $\leq$ *2* $*$ *ereal*(*dist x y*)
⟨*proof*⟩

lemma *Gromov-completion-dist-le-1* [*simp*, *mono-intros*]:
  fixes *x y*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion*
  shows *dist x y* $\leq$ *1*
⟨*proof*⟩

To avoid computations with exponentials, the following lemma is very convenient. It asserts that if x is close enough to infinity, and y is close enough to x, then the Gromov product between x and y is large.

lemma *large-Gromov-product-approx*:
  assumes (*M*::*ereal*) $< \infty$
  shows $\exists$ *e D*. *e* $>$ *0* $\wedge$ *D* $< \infty \wedge$ ($\forall$ *x y*. *dist x y* $\leq$ *e* $\longrightarrow$ *extended-Gromov-distance x* (*to-Gromov-completion basepoint*) $\geq$ *D* $\longrightarrow$ *extended-Gromov-product-at basepoint x y* $\geq$ *M*)
⟨*proof*⟩

On the other hand, far away from infinity, it is equivalent to control the extended Gromov distance or the new distance on the space.

lemma *inside-Gromov-distance-approx*:
  assumes *C* $<$ (∞::*ereal*)
  shows $\exists$ *e* $>$ *0*. $\forall$ *x y*. *extended-Gromov-distance* (*to-Gromov-completion basepoint*) *x* $\leq$ *C* $\longrightarrow$ *dist x y* $\leq$ *e*
    $\longrightarrow$ *esqrt*(*extended-Gromov-distance x y*) $\leq$ *2* $*$ *ereal*(*dist x y*)
⟨*proof*⟩

15.4 Characterizing convergence in the Gromov boundary

The convergence of sequences in the Gromov boundary can be characterized, essentially by definition: sequences tend to a point at infinity iff the Gromov product with this point tends to infinity, while sequences tend to a point inside iff the extended distance tends to 0. In both cases, it is just a matter of unfolding the definition of the distance, and see which one of the two terms (exponential of minus the Gromov product, or extended distance) realizes the minimum. We have constructed the distance essentially so that this property is satisfied.

We could also have defined first the topology, satisfying these conditions, but then we would have had to check that it coincides with the topology that the distance defines, so it seems more economical to proceed in this way.

lemma *Gromov-completion-boundary-limit*:
assumes $x \in$ *Gromov-boundary*
shows $(u \longrightarrow x)$ $F \longleftrightarrow ((\lambda n.$ *extended-Gromov-product-at basepoint* $(u\ n)\ x) \longrightarrow \infty)$ F
⟨*proof*⟩

lemma *extended-Gromov-product-tendsto-PInf-a-b*:
assumes $((\lambda n.$ *extended-Gromov-product-at* $a\ (u\ n)\ (v\ n)) \longrightarrow \infty)$ F
shows $((\lambda n.$ *extended-Gromov-product-at* $b\ (u\ n)\ (v\ n)) \longrightarrow \infty)$ F
⟨*proof*⟩

lemma *Gromov-completion-inside-limit*:
assumes $x \notin$ *Gromov-boundary*
shows $(u \longrightarrow x)$ $F \longleftrightarrow ((\lambda n.$ *extended-Gromov-distance* $(u\ n)\ x) \longrightarrow 0)$ F
⟨*proof*⟩

lemma *to-Gromov-completion-lim* [*simp*, *tendsto-intros*]:
$((\lambda n.$ *to-Gromov-completion* $(u\ n)) \longrightarrow$ *to-Gromov-completion* $a)$ $F \longleftrightarrow (u \longrightarrow a)$ F
⟨*proof*⟩

Now, we can also come back to our original definition of the completion, where points on the boundary correspond to equivalence classes of sequences whose mutual Gromov product tends to infinity. We show that this is compatible with our topology: the sequences that are in the equivalence class of a point on the boundary are exactly the sequences that converge to this point. This is also a direct consequence of the definitions, although the proof requires some unfolding (and playing with the hyperbolicity inequality several times).

First, we show that a sequence in the equivalence class of x converges to x.

lemma *Gromov-completion-converge-to-boundary-aux*:
assumes $x \in$ *Gromov-boundary abs-Gromov-completion* $v = x$ *Gromov-completion-rel* $v\ v$
shows $(\lambda n.$ *extended-Gromov-product-at basepoint* (*to-Gromov-completion* $(v\ n))\ x) \longrightarrow \infty$
⟨*proof*⟩

Then, we prove the converse and therefore the equivalence.

lemma *Gromov-completion-converge-to-boundary*:
assumes $x \in$ *Gromov-boundary*
shows $((\lambda n.$ *to-Gromov-completion* $(u\ n)) \longrightarrow x) \longleftrightarrow$ (*Gromov-completion-rel* $u\ u \wedge$ *abs-Gromov-completion* $u = x)$

⟨*proof*⟩

In particular, it follows that a sequence which is `Gromov_converging_at_boundary` is indeed converging to a point on the boundary, the equivalence class of this sequence.

lemma *Gromov-converging-at-boundary-converges*:
assumes *Gromov-converging-at-boundary u*
shows $\exists x \in$ *Gromov-boundary*. (λn. *to-Gromov-completion* (u n)) $\longrightarrow x$
⟨*proof*⟩

lemma *Gromov-converging-at-boundary-converges'*:
assumes *Gromov-converging-at-boundary u*
shows *convergent* (λn. *to-Gromov-completion* (u n))
⟨*proof*⟩

lemma *lim-imp-Gromov-converging-at-boundary*:
fixes u::*nat* $\Rightarrow$ *'a*::*Gromov-hyperbolic-space*
assumes (λn. *to-Gromov-completion* (u n)) $\longrightarrow x$ $x \in$ *Gromov-boundary*
shows *Gromov-converging-at-boundary u*
⟨*proof*⟩

If two sequences tend to the same point at infinity, then their Gromov product tends to infinity.

lemma *same-limit-imp-Gromov-product-tendsto-infinity*:
assumes $z \in$ *Gromov-boundary*
(λn. *to-Gromov-completion* (u n)) $\longrightarrow z$
(λn. *to-Gromov-completion* (v n)) $\longrightarrow z$
shows $\exists N. \forall n \geq N. \forall m \geq N.$ *Gromov-product-at* a (u n) (v m) $\geq C$
⟨*proof*⟩

An admissible sequence converges in the Gromov boundary, to the point it defines. This follows from the definition of the topology in the two cases, inner and boundary.

lemma *abs-Gromov-completion-limit*:
assumes *Gromov-completion-rel u u*
shows (λn. *to-Gromov-completion* (u n)) $\longrightarrow$ *abs-Gromov-completion u*
⟨*proof*⟩

In particular, a point in the Gromov boundary is the limit of its representative sequence in the space.

lemma *rep-Gromov-completion-limit*:
(λn. *to-Gromov-completion* (*rep-Gromov-completion* x n)) $\longrightarrow x$
⟨*proof*⟩

15.5 Continuity properties of the extended Gromov product and distance

We have defined our extended Gromov product in terms of sequences satisfying the equivalence relation. However, we would like to avoid this definition as much as possible, and express things in terms of the topology of the space. Hence, we reformulate this definition in topological terms, first when one of the two points is inside and the other one is on the boundary, then for all cases, and then we come back to the case where one point is inside, removing the assumption that the other one is on the boundary.

lemma *extended-Gromov-product-inside-boundary-aux*:
assumes *y* ∈ *Gromov-boundary*
shows *extended-Gromov-product-at e* (*to-Gromov-completion x*) *y* = *Inf* {*liminf* (*λn. ereal*(*Gromov-product-at e x* (*v n*))) |*v*. (*λn. to-Gromov-completion* (*v n*)) ⟶ *y*}
⟨*proof*⟩

lemma *extended-Gromov-product-boundary-inside-aux*:
assumes *y* ∈ *Gromov-boundary*
shows *extended-Gromov-product-at e y* (*to-Gromov-completion x*) = *Inf* {*liminf* (*λn. ereal*(*Gromov-product-at e* (*v n*) *x*)) |*v*. (*λn. to-Gromov-completion* (*v n*)) ⟶ *y*}
⟨*proof*⟩

lemma *extended-Gromov-product-at-topological*:
extended-Gromov-product-at e x y = *Inf* {*liminf* (*λn. ereal*(*Gromov-product-at e* (*u n*) (*v n*))) |*u v*. (*λn. to-Gromov-completion* (*u n*)) ⟶ *x* ∧ (*λn. to-Gromov-completion* (*v n*)) ⟶ *y*}
⟨*proof*⟩

lemma *extended-Gromov-product-inside-boundary*:
extended-Gromov-product-at e (*to-Gromov-completion x*) *y* = *Inf* {*liminf* (*λn. ereal*(*Gromov-product-at e x* (*v n*))) |*v*. (*λn. to-Gromov-completion* (*v n*)) ⟶ *y*}
⟨*proof*⟩

lemma *extended-Gromov-product-boundary-inside*:
extended-Gromov-product-at e y (*to-Gromov-completion x*) = *Inf* {*liminf* (*λn. ereal*(*Gromov-product-at e* (*v n*) *x*)) |*v*. (*λn. to-Gromov-completion* (*v n*)) ⟶ *y*}
⟨*proof*⟩

Now, we compare the extended Gromov product to a sequence of Gromov products for converging sequences. As the extended Gromov product is defined as an Inf of limings, it is clearly smaller than the liminf. More interestingly, it is also of the order of magnitude of the limsup, for whatever sequence one uses. In other words, it is canonically defined, up to 2δ.

lemma *extended-Gromov-product-le-liminf*:

assumes ($\lambda n.$ *to-Gromov-completion* (u n)) $\longrightarrow$ *xi*
($\lambda n.$ *to-Gromov-completion* (v n)) $\longrightarrow$ *eta*
shows *liminf* ($\lambda n.$ *Gromov-product-at* e (u n) (v n)) $\geq$ *extended-Gromov-product-at* e *xi eta*
⟨*proof*⟩

lemma *limsup-le-extended-Gromov-product-inside*:
assumes ($\lambda n.$ *to-Gromov-completion* (v n)) $\longrightarrow$ (*eta*::('*a*::*Gromov-hyperbolic-space*) *Gromov-completion*)
shows *limsup* ($\lambda n.$ *Gromov-product-at* e x (v n)) $\leq$ *extended-Gromov-product-at* e (*to-Gromov-completion* x) *eta* + *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

lemma *limsup-le-extended-Gromov-product-inside'*:
assumes ($\lambda n.$ *to-Gromov-completion* (v n)) $\longrightarrow$ (*eta*::('*a*::*Gromov-hyperbolic-space*) *Gromov-completion*)
shows *limsup* ($\lambda n.$ *Gromov-product-at* e (v n) x) $\leq$ *extended-Gromov-product-at* e *eta* (*to-Gromov-completion* x) + *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

lemma *limsup-le-extended-Gromov-product*:
assumes ($\lambda n.$ *to-Gromov-completion* (u n)) $\longrightarrow$ (*xi*::('*a*::*Gromov-hyperbolic-space*) *Gromov-completion*)
($\lambda n.$ *to-Gromov-completion* (v n)) $\longrightarrow$ *eta*
shows *limsup* ($\lambda n.$ *Gromov-product-at* e (u n) (v n)) $\leq$ *extended-Gromov-product-at* e *xi eta* + *2* $*$ *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

One can then extend to the boundary the fact that $(y, z)_x + (x, z)_y = d(x, y)$, up to a constant δ, by taking this identity inside and passing to the limit.

lemma *extended-Gromov-product-add-le*:
extended-Gromov-product-at x *xi* (*to-Gromov-completion* y) + *extended-Gromov-product-at* y *xi* (*to-Gromov-completion* x) $\leq$ *dist* x y
⟨*proof*⟩

lemma *extended-Gromov-product-add-ge*:
extended-Gromov-product-at (x::'*a*::*Gromov-hyperbolic-space*) *xi* (*to-Gromov-completion* y) + *extended-Gromov-product-at* y *xi* (*to-Gromov-completion* x) $\geq$ *dist* x y $-$ *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

If one perturbs a sequence inside the space by a bounded distance, one does not change the limit on the boundary.

lemma *Gromov-converging-at-boundary-bounded-perturbation*:
assumes ($\lambda n.$ *to-Gromov-completion* (u n)) $\longrightarrow$ x
$x \in$ *Gromov-boundary*
$\bigwedge n.$ *dist* (u n) (v n) $\leq C$
shows ($\lambda n.$ *to-Gromov-completion* (v n)) $\longrightarrow$ x
⟨*proof*⟩

We prove that the extended Gromov distance is a continuous function of one variable, by separating the different cases at infinity and inside the space. Note that it is not a continuous function of both variables: if u_n is inside the space but tends to a point x in the boundary, then the extended Gromov distance between u_n and u_n is 0, but for the limit it is ∞.

lemma *extended-Gromov-distance-continuous*:
continuous-on UNIV (λy. *extended-Gromov-distance x y*)
⟨*proof*⟩

lemma *extended-Gromov-distance-continuous'*:
continuous-on UNIV (λx. *extended-Gromov-distance x y*)
⟨*proof*⟩

15.6 Topology of the Gromov boundary

We deduce the basic fact that the original space is open in the Gromov completion from the continuity of the extended distance.

lemma *to-Gromov-completion-range-open*:
open (*range to-Gromov-completion*)
⟨*proof*⟩

lemma *Gromov-boundary-closed*:
closed Gromov-boundary
⟨*proof*⟩

The original space is also dense in its Gromov completion, as all points at infinity are by definition limits of some sequence in the space.

lemma *to-Gromov-completion-range-dense* [*simp*]:
closure (*range to-Gromov-completion*) = *UNIV*
⟨*proof*⟩

lemma *to-Gromov-completion-homeomorphism*:
homeomorphism-on UNIV to-Gromov-completion
⟨*proof*⟩

lemma *to-Gromov-completion-continuous*:
continuous-on UNIV to-Gromov-completion
⟨*proof*⟩

lemma *from-Gromov-completion-continuous*:
homeomorphism-on (*range to-Gromov-completion*) *from-Gromov-completion*
continuous-on (*range to-Gromov-completion*) *from-Gromov-completion*
$\bigwedge x$::('*a*::*Gromov-hyperbolic-space*) *Gromov-completion*. $x \in$ *range to-Gromov-completion*
$\Longrightarrow$ *continuous* (*at x*) *from-Gromov-completion*
⟨*proof*⟩

The Gromov boundary is always complete. Indeed, consider a Cauchy sequence u_n in the boundary, and approximate well enough u_n by a point v_n

inside. Then the sequence v_n is Gromov converging at infinity (the respective Gromov products tend to infinity essentially by definition), and its limit point is the limit of the original sequence u.

proposition *Gromov-boundary-complete*:
complete Gromov-boundary
⟨*proof*⟩

When the initial space is complete, then the whole Gromov completion is also complete: for Cauchy sequences tending to the Gromov boundary, then the convergence is proved as in the completeness of the boundary above. For Cauchy sequences that remain bounded, the convergence is reduced to the convergence inside the original space, which holds by assumption.

proposition *Gromov-completion-complete*:
assumes *complete* (*UNIV*::*'a*::*Gromov-hyperbolic-space set*)
shows *complete* (*UNIV*::*'a Gromov-completion set*)
⟨*proof*⟩

instance *Gromov-completion*::({*Gromov-hyperbolic-space*, *complete-space*}) *complete-space*
⟨*proof*⟩

When the original space is proper, i.e., closed balls are compact, and geodesic, then the Gromov completion (and therefore the Gromov boundary) are compact. The idea to extract a convergent subsequence of a sequence u_n in the boundary is to take the point v_n at distance T along a geodesic tending to the point u_n on the boundary, where T is fixed and large. The points v_n live in a bounded subset of the space, hence they have a convergent subsequence $v_{j(n)}$. It follows that $u_{j(n)}$ is almost converging, up to an error that tends to 0 when T tends to infinity. By a diagonal argument, we obtain a convergent subsequence of u_n.

As we have already proved that the space is complete, there is a shortcut to the above argument, avoiding subsequences and diagonal argument altogether. Indeed, in a complete space it suffices to show that for any $\epsilon > 0$ it is covered by finitely many balls of radius ϵ to get the compactness. This is what we do in the following proof, although the argument is precisely modelled on the first proof we have explained.

theorem *Gromov-completion-compact*:
assumes *proper* (*UNIV*::*'a*::*Gromov-hyperbolic-space-geodesic set*)
shows *compact* (*UNIV*::*'a Gromov-completion set*)
⟨*proof*⟩

If the inner space is second countable, so is its completion, as the former is dense in the latter.

instance *Gromov-completion*::({*Gromov-hyperbolic-space*, *second-countable-topology*}) *second-countable-topology*

⟨*proof*⟩

The same follows readily for the Polish space property.

instance *metric-completion*::({*Gromov-hyperbolic-space*, *polish-space*}) *polish-space*
⟨*proof*⟩

15.7 The Gromov completion of the real line.

We show in the paragraph that the Gromov completion of the real line is obtained by adding one point at $+\infty$ and one point at $-\infty$. In other words, it coincides with ereal.

To show this, we have to understand which sequences of reals are Gromov-converging to the boundary. We show in the next lemma that they are exactly the sequences that converge to $-\infty$ or to $+\infty$.

lemma *real-Gromov-converging-to-boundary*:
fixes $u::nat \Rightarrow real$
shows *Gromov-converging-at-boundary* $u \longleftrightarrow ((u \longrightarrow \infty) \vee (u \longrightarrow -\infty))$
⟨*proof*⟩

There is one single point at infinity in the Gromov completion of reals, i.e., two sequences tending to infinity are equivalent.

lemma *real-Gromov-completion-rel-PInf*:
fixes $u\ v::nat \Rightarrow real$
assumes $u \longrightarrow \infty\ v \longrightarrow \infty$
shows *Gromov-completion-rel* $u\ v$
⟨*proof*⟩

There is one single point at minus infinity in the Gromov completion of reals, i.e., two sequences tending to minus infinity are equivalent.

lemma *real-Gromov-completion-rel-MInf*:
fixes $u\ v::nat \Rightarrow real$
assumes $u \longrightarrow -\infty\ v \longrightarrow -\infty$
shows *Gromov-completion-rel* $u\ v$
⟨*proof*⟩

It follows from the two lemmas above that the Gromov completion of reals is obtained by adding one single point at infinity and one single point at minus infinity. Hence, it is in bijection with the extended reals.

function *to-real-Gromov-completion*::*ereal* $\Rightarrow$ *real Gromov-completion*
where *to-real-Gromov-completion* (*ereal r*) = *to-Gromov-completion r*
| *to-real-Gromov-completion* (∞) = *abs-Gromov-completion* $(\lambda n.\ n)$
| *to-real-Gromov-completion* $(-\infty)$ = *abs-Gromov-completion* $(\lambda n.\ -n)$
⟨*proof*⟩
termination ⟨*proof*⟩

To prove the bijectivity, we prove by hand injectivity and surjectivity using the above lemmas.

lemma *bij-to-real-Gromov-completion*:
bij to-real-Gromov-completion
⟨*proof*⟩

Next, we prove that we have a homeomorphism. By compactness of ereals, it suffices to show that the inclusion map is continuous everywhere. It would be a pain to distinguish all the time if points are at infinity or not, we rather use a criterion saying that it suffices to prove sequential continuity for sequences taking values in a dense subset of the space, here we take the reals. Hence, it suffices to show that if a sequence of reals v_n converges to a limit a in the extended reals, then the image of v_n in the Gromov completion (which is an inner point) converges to the point corresponding to a. We treat separately the cases $a \in \mathbb{R}$, $a = \infty$ and $a = -\infty$. In the first case, everything is trivial. In the other cases, we have characterized in general sequences inside the space that converge to a boundary point, as sequences in the equivalence class defining this boundary point. Since we have described explicitly these equivalence classes in the case of the Gromov completion of the reals (they are respectively the sequences tending to ∞ and to $-\infty$), the result follows readily without any additional computation.

proposition *homeo-to-real-Gromov-completion*:
homeomorphism-on UNIV to-real-Gromov-completion
⟨*proof*⟩

end

theory *Boundary-Extension*
imports *Morse-Gromov-Theorem Gromov-Boundary*
begin

16 Extension of quasi-isometries to the boundary

In this section, we show that a quasi-isometry between geodesic Gromov hyperbolic spaces extends to a homeomorphism between their boundaries.

Applying a quasi-isometry on a geodesic triangle essentially sends it to a geodesic triangle, in hyperbolic spaces. It follows that, up to an additive constant, the Gromov product, which is the distance to the center of the triangle, is multiplied by a constant between λ^{-1} and λ when one applies a quasi-isometry. This argument is given in the next lemma. This implies that two points are close in the Gromov completion if and only if their images are also close in the Gromov completion of the image. Essentially, this lemma implies that a quasi-isometry has a continuous extension to the Gromov boundary, which is a homeomorphism.

lemma *Gromov-product-at-quasi-isometry*:

fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* ⇒ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
shows *Gromov-product-at* (*f x*) (*f y*) (*f z*) ≥ *Gromov-product-at x y z* / *lambda* − *187* ∗ *lambda*^*2* ∗ (*C* + *deltaG*(*TYPE*(*'a*)) + *deltaG*(*TYPE*(*'b*)))
Gromov-product-at (*f x*) (*f y*) (*f z*) ≤ *lambda* ∗ *Gromov-product-at x y z* + *187* ∗ *lambda*^*2* ∗ (*C* + *deltaG*(*TYPE*(*'a*)) + *deltaG*(*TYPE*(*'b*)))
⟨*proof*⟩

lemma *Gromov-converging-at-infinity-quasi-isometry*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* ⇒ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
shows *Gromov-converging-at-boundary* (λ*n*. *f* (*u n*)) ⟷ *Gromov-converging-at-boundary u*
⟨*proof*⟩

We define the extension to the completion of a function $f : X \to Y$ where X and Y are geodesic Gromov-hyperbolic spaces, as a function from $X \cup \partial X$ to $Y \cup \partial Y$, as follows. If x is in the space, we just use $f(x)$ (with the suitable coercions for the definition). Otherwise, we wish to define $f(x)$ as the limit of $f(u_n)$ for all sequences tending to x. For the definition, we use one such sequence chosen arbitrarily (this is the role of `rep_Gromov_completion x` below, it is indeed a sequence in the space tending to x), and we use the limit of $f(u_n)$ (if it exists, otherwise the framework will choose some point for us but it will make no sense whatsoever).

For quasi-isometries, we have indeed that $f(u_n)$ converges if u_n converges to a boundary point, by `Gromov_converging_at_infinity_quasi_isometry`, so this definition is meaningful. Moreover, continuity of the extension follows readily from this (modulo a suitable criterion for continuity based on sequences convergence, established in `continuous_at_extension_sequentially'`).

definition *Gromov-extension*::(*'a*::*Gromov-hyperbolic-space* ⇒ *'b*::*Gromov-hyperbolic-space*) ⇒ (*'a Gromov-completion* ⇒ *'b Gromov-completion*)
where *Gromov-extension f x* = (*if x* ∈ *Gromov-boundary then lim* (*to-Gromov-completion o f o* (*rep-Gromov-completion x*))
else to-Gromov-completion (*f* (*from-Gromov-completion x*)))

lemma *Gromov-extension-inside-space* [*simp*]:
Gromov-extension f (*to-Gromov-completion x*) = *to-Gromov-completion* (*f x*)
⟨*proof*⟩

lemma *Gromov-extension-id* [*simp*]:
Gromov-extension (*id*::*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) = *id*
Gromov-extension (λ*x*::*'a*. *x*) = (λ*x*. *x*)
⟨*proof*⟩

The Gromov extension of a quasi-isometric map sends the boundary to the boundary.

lemma *Gromov-extension-quasi-isometry-boundary-to-boundary*:
fixes f::$'a$::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ $'b$::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
$x \in$ *Gromov-boundary*
shows (*Gromov-extension f*) $x \in$ *Gromov-boundary*
⟨*proof*⟩

If the original function is continuous somewhere inside the space, then its Gromov extension is continuous at the corresponding point inside the completion. This is clear as the original space is open in the Gromov completion, but the proof requires to go back and forth between one space and the other.

lemma *Gromov-extension-continuous-inside*:
fixes f::$'a$::*Gromov-hyperbolic-space* $\Rightarrow$ $'b$::*Gromov-hyperbolic-space*
assumes *continuous* (*at x within S*) *f*
shows *continuous* (*at* (*to-Gromov-completion x*) *within* (*to-Gromov-completion*‘*S*)) (*Gromov-extension f*)
⟨*proof*⟩

The extension to the boundary of a quasi-isometry is continuous. This is a nontrivial statement, but it follows readily from the fact we have already proved that sequences converging at the boundary are mapped to sequences converging to the boundary. The proof is expressed using a convenient continuity criterion for which we only need to control what happens for sequences inside the space.

proposition *Gromov-extension-continuous*:
fixes f::$'a$::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ $'b$::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
$x \in$ *Gromov-boundary*
shows *continuous* (*at x*) (*Gromov-extension f*)
⟨*proof*⟩

Combining the two previous statements on continuity inside the space and continuity at the boundary, we deduce that a continuous quasi-isometry extends to a continuous map everywhere.

proposition *Gromov-extension-continuous-everywhere*:
fixes f::$'a$::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ $'b$::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
continuous-on UNIV f
shows *continuous-on UNIV* (*Gromov-extension f*)
⟨*proof*⟩

The extension to the boundary is functorial on the category of quasi-isometries, i.e., the composition of extensions is the extension of the composition. This is clear inside the space, and it follows from the continuity at boundary points.

lemma *Gromov-extension-composition*:
fixes f::$'a$::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ $'b$::*Gromov-hyperbolic-space-geodesic*

and *g*::*'b*::*Gromov-hyperbolic-space-geodesic* ⇒ *'c*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
mu D−quasi-isometry g
shows *Gromov-extension* (*g o f*) = *Gromov-extension g o Gromov-extension f*
⟨*proof*⟩

Now, we turn to the same kind of statement, but for homeomorphisms. We claim that if a quasi-isometry f is a homeomorphism on a subset X of the space, then its extension is a homeomorphism on X union the boundary of the space. For the proof, we have to show that a sequence u_n tends to a point x if and only if $f(u_n)$ tends to $f(x)$. We separate the cases x in the boundary, and x inside the space. For x in the boundary, we use a homeomorphism criterion expressed solely in terms of sequences converging to the boundary, for which we already know everything. For x in the space, the proof is straightforward, but tedious. We argue that eventually u_n is in the space for the direct implication, or $f(u_n)$ is in the space for the second implication, and then we use that f is a homeomorphism inside the space to conclude.

lemma *Gromov-extension-homeomorphism*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* ⇒ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
homeomorphism-on X f
shows *homeomorphism-on* (*to-Gromov-completion'X* ∪ *Gromov-boundary*) (*Gromov-extension f*)
⟨*proof*⟩

In particular, it follows that the extension to the boundary of a quasi-isometry is always a homeomorphism, regardless of the continuity properties of the original map.

proposition *Gromov-extension-boundary-homeomorphism*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* ⇒ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry f*
shows *homeomorphism-on Gromov-boundary* (*Gromov-extension f*)
⟨*proof*⟩

When the quasi-isometric embedding is a quasi-isometric isomorphism, i.e., it is onto up to a bounded distance C, then its Gromov extension is onto on the boundary. Indeed, a point in the image boundary is a limit of a sequence inside the space. Perturbing by a bounded distance (which does not change the asymptotic behavior), it is the limit of a sequence inside the image of f. Then the preimage under f of this sequence does converge, and its limit is sent by the extension on the original point, proving the surjectivity.

lemma *Gromov-extension-onto*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* ⇒ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry-between UNIV UNIV f*
y ∈ *Gromov-boundary*

shows $\exists\, x \in$ *Gromov-boundary. Gromov-extension f x* $=$ *y*
⟨*proof*⟩

lemma *Gromov-extension-onto'*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry-between UNIV UNIV f*
shows (*Gromov-extension f*)*‘Gromov-boundary* $=$ *Gromov-boundary*
⟨*proof*⟩

Finally, we obtain that a quasi-isometry between two Gromov hyperbolic spaces induces a homeomorphism of their boundaries.

theorem *Gromov-boundaries-homeomorphic*:
fixes *f*::*'a*::*Gromov-hyperbolic-space-geodesic* $\Rightarrow$ *'b*::*Gromov-hyperbolic-space-geodesic*
assumes *lambda C−quasi-isometry-between UNIV UNIV f*
shows (*Gromov-boundary*::*'a Gromov-completion set*) *homeomorphic* (*Gromov-boundary*::*'b Gromov-completion set*)
⟨*proof*⟩

17 Extensions of isometries to the boundary

The results of the previous section can be improved for isometries, as there is no need for geodesicity any more. We follow the same proofs as in the previous section

An isometry preserves the Gromov product.

lemma *Gromov-product-isometry*:
assumes *isometry-on UNIV f*
shows *Gromov-product-at* (*f x*) (*f y*) (*f z*) $=$ *Gromov-product-at x y z*
⟨*proof*⟩

An isometry preserves convergence at infinity.

lemma *Gromov-converging-at-infinity-isometry*:
fixes *f*::*'a*::*Gromov-hyperbolic-space* $\Rightarrow$ *'b*::*Gromov-hyperbolic-space*
assumes *isometry-on UNIV f*
shows *Gromov-converging-at-boundary* ($\lambda n.\ f\ (u\ n)$) $\longleftrightarrow$ *Gromov-converging-at-boundary u*
⟨*proof*⟩

The Gromov extension of an isometry sends the boundary to the boundary.

lemma *Gromov-extension-isometry-boundary-to-boundary*:
fixes *f*::*'a*::*Gromov-hyperbolic-space* $\Rightarrow$ *'b*::*Gromov-hyperbolic-space*
assumes *isometry-on UNIV f*
$x \in$ *Gromov-boundary*
shows (*Gromov-extension f*) $x \in$ *Gromov-boundary*
⟨*proof*⟩

The Gromov extension of an isometry is a homeomorphism. (We copy the proof for quasi-isometries, with some simplifications.)

lemma *Gromov-extension-isometry-homeomorphism*:
  fixes *f*::*'a*::*Gromov-hyperbolic-space* ⇒ *'b*::*Gromov-hyperbolic-space*
  assumes *isometry-on UNIV f*
  shows *homeomorphism-on UNIV* (*Gromov-extension f*)
⟨*proof*⟩

The composition of the Gromov extension of two isometries is the Gromov extension of the composition.

lemma *Gromov-extension-isometry-on-composition*:
  assumes *isometry-on UNIV f*
    isometry-on UNIV g
  shows *Gromov-extension* (*g o f*) = *Gromov-extension g o Gromov-extension f*
⟨*proof*⟩

We specialize the previous results to bijective isometries, as this is the setting where they will be used most of the time.

lemma *Gromov-extension-isometry*:
  assumes *isometry f*
  shows *homeomorphism-on UNIV* (*Gromov-extension f*)
    continuous-on UNIV (*Gromov-extension f*)
    continuous (*at x*) (*Gromov-extension f*)
⟨*proof*⟩

lemma *Gromov-extension-isometry-composition*:
  assumes *isometry f*
    isometry g
  shows *Gromov-extension* (*g o f*) = *Gromov-extension g o Gromov-extension f*
⟨*proof*⟩

lemma *Gromov-extension-isometry-iterates*:
  fixes *f*::*'a* ⇒ (*'a*::*Gromov-hyperbolic-space*)
  assumes *isometry f*
  shows *Gromov-extension* (*f^^n*) = (*Gromov-extension f*)*^^n*
⟨*proof*⟩

lemma *Gromov-extension-isometry-inv*:
  assumes *isometry f*
  shows *inv* (*Gromov-extension f*) = *Gromov-extension* (*inv f*)
    bij (*Gromov-extension f*)
⟨*proof*⟩

We will especially use fixed points on the boundary. We note that if a point is fixed by (the Gromov extension of) a map, then it is fixed by (the Gromov extension of) its inverse.

lemma *Gromov-extension-inv-fixed-point*:
  assumes *isometry* (*f*::*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) *Gromov-extension f xi* = *xi*
  shows *Gromov-extension* (*inv f*) *xi* = *xi*

⟨*proof*⟩

The extended Gromov product is invariant under isometries. This follows readily from the definition, but still the proof is not fully automatic, unfortunately.

lemma *Gromov-extension-preserves-extended-Gromov-product*:
assumes *isometry f*
shows *extended-Gromov-product-at* (*f x*) (*Gromov-extension f xi*) (*Gromov-extension f eta*) = *extended-Gromov-product-at x xi eta*
⟨*proof*⟩

end

18 Busemann functions

theory *Busemann-Function*
imports *Boundary-Extension Ergodic-Theory.Fekete*
begin

The Busemann function $B_\xi(x, y)$ measures the difference $d(\xi, x) - d(\xi, y)$, where ξ is a point at infinity and x and y are inside a Gromov hyperbolic space. This is not well defined in this way, as we are subtracting two infinities, but one can make sense of this difference by considering the behavior along a sequence tending to ξ. The limit may depend on the sequence, but as usual in Gromov hyperbolic spaces it only depends on the sequence up to a uniform constant. Thus, we may define the Busemann function using for instance the supremum of the limsup over all possible sequences – other choices would give rise to equivalent definitions, up to some multiple of δ.

definition *Busemann-function-at*::(*'a*::*Gromov-hyperbolic-space*) *Gromov-completion* ⇒ *'a* ⇒ *'a* ⇒ *real*
where *Busemann-function-at xi x y* = *real-of-ereal* (
Sup {*limsup* (*λn. ereal*(*dist x* (*u n*) − *dist y* (*u n*))) |*u*. (*λn. to-Gromov-completion* (*u n*)) ⟶ *xi*})

Since limsups are only defined for complete orders currently, the definition goes through ereals, and we go back to reals afterwards. However, there is no real difficulty here, as eveything is bounded above and below (by $d(x, y)$ and $-d(x, y)$ respectively.

lemma *Busemann-function-ereal*:
ereal(*Busemann-function-at xi x y*) = *Sup* {*limsup* (*λn. ereal*(*dist x* (*u n*) − *dist y* (*u n*))) |*u*. (*λn. to-Gromov-completion* (*u n*)) ⟶ *xi*}
⟨*proof*⟩

If ξ is not at infinity, then the Busemann function is simply the difference of the distances.

lemma *Busemann-function-inner*:

Busemann-function-at (*to-Gromov-completion z*) *x y* = *dist x z* − *dist y z*
⟨*proof*⟩

The Busemann function measured at the same points vanishes.

lemma *Busemann-function-xx* [*simp*]:
Busemann-function-at xi x x = *0*
⟨*proof*⟩

Perturbing the points gives rise to a variation of the Busemann function bounded by the size of the variations. This is obvious for inner Busemann functions, and everything passes readily to the limit.

lemma *Busemann-function-mono* [*mono-intros*]:
Busemann-function-at xi x y ≤ *Busemann-function-at xi x' y'* + *dist x x'* + *dist y y'*
⟨*proof*⟩

In particular, it follows that the Busemann function $B_\xi(x, y)$ is bounded in absolute value by $d(x, y)$.

lemma *Busemann-function-le-dist* [*mono-intros*]:
abs(*Busemann-function-at xi x y*) ≤ *dist x y*
⟨*proof*⟩

lemma *Busemann-function-Lipschitz* [*mono-intros*]:
abs(*Busemann-function-at xi x y* − *Busemann-function-at xi x' y'*) ≤ *dist x x'* + *dist y y'*
⟨*proof*⟩

By the very definition of the Busemann function, the difference of distance functions is bounded above by the Busemann function when one converges to ξ.

lemma *Busemann-function-limsup*:
assumes (λ*n*. *to-Gromov-completion* (*u n*)) ⟶ *xi*
shows *limsup* (λ*n*. *dist x* (*u n*) − *dist y* (*u n*)) ≤ *Busemann-function-at xi x y*
⟨*proof*⟩

There is also a corresponding bound below, but with the loss of a constant. This follows from the hyperbolicity of the space and a simple computation.

lemma *Busemann-function-liminf*:
assumes (λ*n*. *to-Gromov-completion* (*u n*)) ⟶ *xi*
shows *Busemann-function-at xi x y* ≤ *liminf* (λ*n*. *dist* (*x*::'*a*::*Gromov-hyperbolic-space*) (*u n*) − *dist y* (*u n*)) + *2* ∗ *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

To avoid formulating things in terms of liminf and limsup on ereal, the following formulation of the two previous lemmas is useful.

lemma *Busemann-function-inside-approx*:

assumes $e > (0::real)$ $(\lambda n.\ to\text{-}Gromov\text{-}completion\ (t\ n::'a::Gromov\text{-}hyperbolic\text{-}space))$ $\longrightarrow xi$
shows $eventually\ (\lambda n.\ Busemann\text{-}function\text{-}at\ (to\text{-}Gromov\text{-}completion\ (t\ n))\ x\ y \leq Busemann\text{-}function\text{-}at\ xi\ x\ y + e$
$\wedge\ Busemann\text{-}function\text{-}at\ (to\text{-}Gromov\text{-}completion\ (t\ n))\ x\ y \geq Busemann\text{-}function\text{-}at\ xi\ x\ y - 2 * deltaG(TYPE('a)) - e)\ sequentially$
⟨*proof*⟩

The Busemann function is essentially a morphism, i.e., it should satisfy $B_\xi(x,z) = B_\xi(x,y) + B_\xi(y,z)$, as it is defined as a difference of distances. This is not exactly the case as there is a choice in the definition, but it is the case up to a uniform constant, as we show in the next few lemmas. One says that it is a *quasi-morphism.*

lemma *Busemann-function-triangle* [*mono-intros*]:
$Busemann\text{-}function\text{-}at\ xi\ x\ z \leq Busemann\text{-}function\text{-}at\ xi\ x\ y + Busemann\text{-}function\text{-}at\ xi\ y\ z$
⟨*proof*⟩

lemma *Busemann-function-xy-yx* [*mono-intros*]:
$Busemann\text{-}function\text{-}at\ xi\ x\ y + Busemann\text{-}function\text{-}at\ xi\ y\ (x::'a::Gromov\text{-}hyperbolic\text{-}space) \leq 2 * deltaG(TYPE('a))$
⟨*proof*⟩

theorem *Busemann-function-quasi-morphism* [*mono-intros*]:
$|Busemann\text{-}function\text{-}at\ xi\ x\ y + Busemann\text{-}function\text{-}at\ xi\ y\ z - Busemann\text{-}function\text{-}at\ xi\ x\ (z::'a::Gromov\text{-}hyperbolic\text{-}space)| \leq 2 * deltaG(TYPE('a))$
⟨*proof*⟩

The extended Gromov product can be bounded from below by the Busemann function.

lemma *Busemann-function-le-Gromov-product*:
$- Busemann\text{-}function\text{-}at\ xi\ y\ x/2 \leq extended\text{-}Gromov\text{-}product\text{-}at\ x\ xi\ (to\text{-}Gromov\text{-}completion\ y)$
⟨*proof*⟩

It follows that, if the Busemann function tends to minus infinity, i.e., the distance to ξ becomes smaller and smaller in a suitable sense, then the sequence is converging to ξ. This is only an implication: one can have sequences tending to ξ for which the Busemann function does not tend to $-\infty$. This is in fact a stronger notion of convergence, sometimes called radial convergence.

proposition *Busemann-function-minus-infinity-imp-convergent*:
assumes $((\lambda n.\ Busemann\text{-}function\text{-}at\ xi\ (u\ n)\ x) \longrightarrow -\infty)\ F$
shows $((\lambda n.\ to\text{-}Gromov\text{-}completion\ (u\ n)) \longrightarrow xi)\ F$
⟨*proof*⟩

Busemann functions are invariant under isometries. This is trivial as everything is defined in terms of the distance, but the definition in terms of supremum and limsups makes the proof tedious.

lemma *Busemann-function-isometry*:
 assumes *isometry f*
 shows *Busemann-function-at* (*Gromov-extension f xi*) (*f x*) (*f y*) = *Busemann-function-at xi x y*
⟨*proof*⟩

lemma *dist-le-max-Busemann-functions* [*mono-intros*]:
 assumes *xi* ≠ *eta*
 shows *dist x* (*y*::′*a*::*Gromov-hyperbolic-space*) ≤ *2* ∗ *real-of-ereal* (*extended-Gromov-product-at y xi eta*)
 + *max* (*Busemann-function-at xi x y*) (*Busemann-function-at eta x y*) + *2* ∗ *deltaG*(*TYPE*(′*a*))
⟨*proof*⟩

lemma *dist-minus-Busemann-max-ineq*:
 dist (*x*::′*a*::*Gromov-hyperbolic-space*) *z* − *Busemann-function-at xi z x* ≤ *max* (*dist x y* − *Busemann-function-at xi y x*) (*dist y z* − *Busemann-function-at xi z y* − *2* ∗ *Busemann-function-at xi y x*) + *8* ∗ *deltaG*(*TYPE*(′*a*))
⟨*proof*⟩

end

19 Classification of isometries on a Gromov hyperbolic space

theory *Isometries-Classification*
 imports *Gromov-Boundary Busemann-Function*
begin

Isometries of Gromov hyperbolic spaces are of three types:

- Elliptic ones, for which orbits are bounded.
- Parabolic ones, which are not elliptic and have exactly one fixed point at infinity.
- Loxodromic ones, which are not elliptic and have exactly two fixed points at infinity.

In this file, we show that all isometries are indeed of this form, and give further properties for each type.

For the definition, we use another characterization in terms of stable translation length: for isometries which are not elliptic, then they are parabolic if the stable translation length is 0, loxodromic if it is positive. This gives a very efficient definition, and it is clear from this definition that the three categories of isometries are disjoint. All the work is then to go from this general definition to the dynamical properties in terms of fixed points on the boundary.

19.1 The translation length

The translation length is the minimal translation distance of an isometry. The stable translation length is the limit of the translation length of f^n divided by n.

definition *translation-length*::((*'a*::*metric-space*) ⇒ *'a*) ⇒ *real*
 where *translation-length f* = *Inf* {*dist x* (*f x*)|*x*. *True*}

lemma *translation-length-nonneg* [*simp*, *mono-intros*]:
 translation-length f ≥ *0*
⟨*proof*⟩

lemma *translation-length-le* [*mono-intros*]:
 translation-length f ≤ *dist x* (*f x*)
⟨*proof*⟩

definition *stable-translation-length*::((*'a*::*metric-space*) ⇒ *'a*) ⇒ *real*
 where *stable-translation-length f* = *Inf* {*translation-length* (*f^^n*)/*n* |*n*. *n* > *0*}

lemma *stable-translation-length-nonneg* [*simp*]:
 stable-translation-length f ≥ *0*
⟨*proof*⟩

lemma *stable-translation-length-le-translation-length* [*mono-intros*]:
 n ∗ *stable-translation-length f* ≤ *translation-length* (*f^^n*)
⟨*proof*⟩

lemma *semicontraction-iterates*:
 fixes *f*::(*'a*::*metric-space*) ⇒ *'a*
 assumes *1*−*lipschitz-on UNIV f*
 shows *1*−*lipschitz-on UNIV* (*f^^n*)
⟨*proof*⟩

If f is a semicontraction, then its stable translation length is the limit of $d(x, f^n x)/n$ for any n. While it is obvious that the liminf of this quantity is at least the stable translation length (which is defined as an inf over all points and all times), the opposite inequality is more interesting. One may find a point y and a time k for which $d(y, f^k y)/k$ is very close to the stable translation length. By subadditivity of the sequence $n \mapsto f(y, f^n y)$ and Fekete's Lemma, it follows that, for any large n, then $d(y, f^n y)/n$ is also very close to the stable translation length. Since this is equal to $d(x, f^n x)/n$ up to $\pm 2d(x, y)/n$, the result follows.

proposition *stable-translation-length-as-pointwise-limit*:
 assumes *1*−*lipschitz-on UNIV f*

shows $(\lambda n.\ dist\ x\ ((f\widehat{\ }\widehat{\ }n)\ x)/n) \longrightarrow$ *stable-translation-length f*
⟨*proof*⟩

It follows from the previous proposition that the stable translation length is also the limit of the renormalized translation length of f^n.

proposition *stable-translation-length-as-limit*:
assumes *1−lipschitz-on UNIV f*
shows $(\lambda n.\ translation\text{-}length\ (f\widehat{\ }\widehat{\ }n)\ /\ n) \longrightarrow$ *stable-translation-length f*
⟨*proof*⟩

lemma *stable-translation-length-inv*:
assumes *isometry f*
shows *stable-translation-length* (*inv f*) = *stable-translation-length f*
⟨*proof*⟩

19.2 The strength of an isometry at a fixed point at infinity

The additive strength of an isometry at a fixed point at infinity is the asymptotic average every point is moved towards the fixed point at each step. It is measured using the Busemann function.

definition *additive-strength*::('*a*::*Gromov-hyperbolic-space* ⇒ '*a*) ⇒ ('*a Gromov-completion*) ⇒ *real*
where *additive-strength f xi* = *lim* (λ*n*. (*Busemann-function-at xi* ((*f*^^*n*) *basepoint*) *basepoint*)/*n*)

For additivity reasons, as the Busemann function is a quasi-morphism, the additive strength measures the deplacement even at finite times. It is also uniform in terms of the basepoint. This shows that an isometry sends horoballs centered at a fixed point to horoballs, up to a uniformly bounded error depending only on δ.

lemma *Busemann-function-eq-additive-strength*:
assumes *isometry f Gromov-extension f xi* = *xi*
shows |*Busemann-function-at xi* ((*f*^^*n*) *x*) (*x*::'*a*::*Gromov-hyperbolic-space*) − *real n* ∗ *additive-strength f xi*| ≤ *2* ∗ *deltaG*(*TYPE*('*a*))
⟨*proof*⟩

lemma *additive-strength-as-limit* [*tendsto-intros*]:
assumes *isometry f Gromov-extension f xi* = *xi*
shows (λ*n*. *Busemann-function-at xi* ((*f*^^*n*) *x*) *x*/*n*) ⟶ *additive-strength f xi*
⟨*proof*⟩

The additive strength measures the amount of displacement towards a fixed point at infinity. Therefore, the distance from x to $f^n x$ is at least n times the additive strength, but one might think that it might be larger, if there is displacement along the horospheres. It turns out that this is not the case:

the displacement along the horospheres is at most logarithmic (this is a classical property of parabolic isometries in hyperbolic spaces), and in fact it is bounded for loxodromic elements. We prove here that the growth is at most logarithmic in all cases, using a small computation based on the hyperbolicity inequality, expressed in Lemma `dist_minus_Busemann_max_ineq` above. This lemma will be used below to show that the translation length is the absolute value of the additive strength.

lemma *dist-le-additive-strength*:
assumes *isometry* (f::$'a$::*Gromov-hyperbolic-space* ⇒ $'a$) *Gromov-extension f xi* = *xi additive-strength f xi* ≥ *0 n* ≥ *1*
shows *dist x* ((*f*^^*n*) *x*) ≤ *dist x* (*f x*) + *real n* ∗ *additive-strength f xi* + *ceiling* (*log 2 n*) ∗ *16* ∗ *deltaG*(*TYPE*($'a$))
⟨*proof*⟩

The strength of the inverse of a map is the opposite of the strength of the map.

lemma *additive-strength-inv*:
assumes *isometry* (f::$'a$::*Gromov-hyperbolic-space* ⇒ $'a$) *Gromov-extension f xi* = *xi*
shows *additive-strength* (*inv f*) *xi* = − *additive-strength f xi*
⟨*proof*⟩

We will now prove that the stable translation length of an isometry is given by the absolute value of its strength at any fixed point. We start with the case where the strength is nonnegative, and then reduce to this case by considering the map or its inverse.

lemma *stable-translation-length-eq-additive-strength-aux*:
assumes *isometry* (f::$'a$::*Gromov-hyperbolic-space* ⇒ $'a$) *Gromov-extension f xi* = *xi additive-strength f xi* ≥ *0*
shows *stable-translation-length f* = *additive-strength f xi*
⟨*proof*⟩

lemma *stable-translation-length-eq-additive-strength*:
assumes *isometry* (f::$'a$::*Gromov-hyperbolic-space* ⇒ $'a$) *Gromov-extension f xi* = *xi*
shows *stable-translation-length f* = *abs*(*additive-strength f xi*)
⟨*proof*⟩

19.3 Elliptic isometries

Elliptic isometries are the simplest ones: they have bounded orbits.

definition *elliptic-isometry*::($'a$::*Gromov-hyperbolic-space* ⇒ $'a$) ⇒ *bool*
where *elliptic-isometry f* = (*isometry f* ∧ (∀ *x*. *bounded* {(*f*^^*n*) *x*|*n*. *True*}))

lemma *elliptic-isometryD*:
assumes *elliptic-isometry f*

shows *bounded* {(*f*^^*n*) *x* |*n*. *True*}
isometry f
⟨*proof*⟩

lemma *elliptic-isometryI* [*intro*]:
assumes *bounded* {(*f*^^*n*) *x* |*n*. *True*}
isometry f
shows *elliptic-isometry f*
⟨*proof*⟩

The inverse of an elliptic isometry is an elliptic isometry.

lemma *elliptic-isometry-inv*:
assumes *elliptic-isometry f*
shows *elliptic-isometry* (*inv f*)
⟨*proof*⟩

The inverse of a bijective map is an elliptic isometry if and only if the original map is.

lemma *elliptic-isometry-inv-iff*:
assumes *bij f*
shows *elliptic-isometry* (*inv f*) ⟷ *elliptic-isometry f*
⟨*proof*⟩

The identity is an elliptic isometry.

lemma *elliptic-isometry-id*:
elliptic-isometry id
⟨*proof*⟩

The translation length of an elliptic isometry is 0.

lemma *elliptic-isometry-stable-translation-length*:
assumes *elliptic-isometry f*
shows *stable-translation-length f* = *0*
⟨*proof*⟩

If an isometry has a fixed point, then it is elliptic.

lemma *isometry-with-fixed-point-is-elliptic*:
assumes *isometry f f x* = *x*
shows *elliptic-isometry f*
⟨*proof*⟩

19.4 Parabolic and loxodromic isometries

An isometry is parabolic if it is not elliptic and if its translation length vanishes.

definition *parabolic-isometry*::('*a*::*Gromov-hyperbolic-space* ⇒ '*a*) ⇒ *bool*
where *parabolic-isometry f* = (*isometry f* ∧ ¬*elliptic-isometry f* ∧ *stable-translation-length f* = *0*)

An isometry is loxodromic if it is not elliptic and if its translation length is nonzero.

definition *loxodromic-isometry*::(*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) ⇒ *bool*
where *loxodromic-isometry f* = (*isometry f* ∧ ¬*elliptic-isometry f* ∧ *stable-translation-length f* ≠ *0*)

The main features of such isometries are expressed in terms of their fixed points at infinity. We define them now, but proving that the definitions make sense will take some work.

definition *neutral-fixed-point*::(*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) ⇒ *'a Gromov-completion*
where *neutral-fixed-point f* = (*SOME xi. xi* ∈ *Gromov-boundary* ∧ *Gromov-extension f xi* = *xi* ∧ *additive-strength f xi* = *0*)

definition *attracting-fixed-point*::(*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) ⇒ *'a Gromov-completion*
where *attracting-fixed-point f* = (*SOME xi. xi* ∈ *Gromov-boundary* ∧ *Gromov-extension f xi* = *xi* ∧ *additive-strength f xi* < *0*)

definition *repelling-fixed-point*::(*'a*::*Gromov-hyperbolic-space* ⇒ *'a*) ⇒ *'a Gromov-completion*
where *repelling-fixed-point f* = (*SOME xi. xi* ∈ *Gromov-boundary* ∧ *Gromov-extension f xi* = *xi* ∧ *additive-strength f xi* > *0*)

lemma *parabolic-isometryD*:
assumes *parabolic-isometry f*
shows *isometry f*
¬*bounded* {(*f*^^*n*) *x*|*n*. *True*}
stable-translation-length f = *0*
¬*elliptic-isometry f*
⟨*proof*⟩

lemma *parabolic-isometryI*:
assumes *isometry f*
¬*bounded* {(*f*^^*n*) *x*|*n*. *True*}
stable-translation-length f = *0*
shows *parabolic-isometry f*
⟨*proof*⟩

lemma *loxodromic-isometryD*:
assumes *loxodromic-isometry f*
shows *isometry f*
¬*bounded* {(*f*^^*n*) *x*|*n*. *True*}
stable-translation-length f > *0*
¬*elliptic-isometry f*
⟨*proof*⟩

To have a loxodromic isometry, it suffices to know that the stable translation length is nonzero, as elliptic isometries have zero translation length.

lemma *loxodromic-isometryI*:
assumes *isometry f*
stable-translation-length f $\neq$ *0*
shows *loxodromic-isometry f*
⟨*proof*⟩

Any isometry is elliptic, or parabolic, or loxodromic, and these possibilities are mutually exclusive.

lemma *elliptic-or-parabolic-or-loxodromic*:
assumes *isometry f*
shows *elliptic-isometry f* $\vee$ *parabolic-isometry f* $\vee$ *loxodromic-isometry f*
⟨*proof*⟩

lemma *elliptic-imp-not-parabolic-loxodromic*:
assumes *elliptic-isometry f*
shows ¬*parabolic-isometry f*
¬*loxodromic-isometry f*
⟨*proof*⟩

lemma *parabolic-imp-not-elliptic-loxodromic*:
assumes *parabolic-isometry f*
shows ¬*elliptic-isometry f*
¬*loxodromic-isometry f*
⟨*proof*⟩

lemma *loxodromic-imp-not-elliptic-parabolic*:
assumes *loxodromic-isometry f*
shows ¬*elliptic-isometry f*
¬*parabolic-isometry f*
⟨*proof*⟩

The inverse of a parabolic isometry is parabolic.

lemma *parabolic-isometry-inv*:
assumes *parabolic-isometry f*
shows *parabolic-isometry* (*inv f*)
⟨*proof*⟩

The inverse of a loxodromic isometry is loxodromic.

lemma *loxodromic-isometry-inv*:
assumes *loxodromic-isometry f*
shows *loxodromic-isometry* (*inv f*)
⟨*proof*⟩

We will now prove that an isometry which is not elliptic has a fixed point at infinity. This is very easy if the space is proper (ensuring that the Gromov completion is compact), but in fact this holds in general. One constructs it by considering a sequence r_n such that $f^{r_n} 0$ tends to infinity, and additionally $d(f^l 0, 0) < d(f^{r_n} 0, 0)$ for $l < r_n$: this implies the convergence at infinity of $f^{r_n} 0$, by an argument based on a Gromov product computation – and the limit is a fixed point. Moreover, it has nonpositive additive strength, essentially by construction.

lemma *high-scores*:
  fixes $u::nat \Rightarrow real$ **and** $i::nat$ **and** $C::real$
  assumes $\neg(bdd\text{-}above\ (range\ u))$
  shows $\exists\, n.\ (\forall\, l \leq n.\ u\ l \leq u\ n) \wedge u\ n \geq C \wedge n \geq i$
⟨*proof*⟩

lemma *isometry-not-elliptic-has-attracting-fixed-point*:
  assumes *isometry f*
    ¬(*elliptic-isometry f*)
  shows ∃ *xi* ∈ *Gromov-boundary*. *Gromov-extension f xi* = *xi* ∧ *additive-strength f xi* ≤ *0*
⟨*proof*⟩

Applying the previous result to the inverse map, we deduce that there is also a fixed point with nonnegative strength.

lemma *isometry-not-elliptic-has-repelling-fixed-point*:
  assumes *isometry f*
    ¬(*elliptic-isometry f*)
  shows ∃ *xi* ∈ *Gromov-boundary*. *Gromov-extension f xi* = *xi* ∧ *additive-strength f xi* ≥ *0*
⟨*proof*⟩

19.4.1 Parabolic isometries

We show that a parabolic isometry has (at least) one neutral fixed point at infinity.

lemma *parabolic-fixed-point*:
  assumes *parabolic-isometry f*
  shows *neutral-fixed-point f* ∈ *Gromov-boundary*
    Gromov-extension f (*neutral-fixed-point f*) = *neutral-fixed-point f*
    additive-strength f (*neutral-fixed-point f*) = *0*
⟨*proof*⟩

Parabolic isometries have exactly one fixed point, the neutral fixed point at infinity. The proof goes as follows: if it has another fixed point, then the orbit of a basepoint would stay on the horospheres centered at both fixed points. But the intersection of two horospheres based at different points is a a bounded set. Hence, the map has a bounded orbit, and is therefore elliptic.

theorem *parabolic-unique-fixed-point*:
  assumes *parabolic-isometry f*
  shows *Gromov-extension f xi* = *xi* ⟷ *xi* = *neutral-fixed-point f*
⟨*proof*⟩

When one iterates a parabolic isometry, the distance to the starting point can grow at most logarithmically.

lemma *parabolic-logarithmic-growth*:
  assumes *parabolic-isometry* $(f::'a::Gromov\text{-}hyperbolic\text{-}space \Rightarrow 'a)$ $n \geq 1$
  shows $dist\ x\ ((f\widehat{\ }\widehat{\ }n)\ x) \leq dist\ x\ (f\ x) + ceiling\ (log\ 2\ n) * 16 * deltaG(TYPE('a))$
$\langle proof \rangle$

It follows that there is no parabolic isometry in trees, since the formula in the previous lemma shows that there is no orbit growth as $\delta = 0$, and therefore orbits are bounded, contradicting the parabolicity of the isometry.

lemma *tree-no-parabolic-isometry*:
  assumes *isometry* $(f::'a::Gromov\text{-}hyperbolic\text{-}space\text{-}0 \Rightarrow 'a)$
  shows *elliptic-isometry f* $\vee$ *loxodromic-isometry f*
$\langle proof \rangle$

19.4.2 Loxodromic isometries

A loxodromic isometry has (at least) two fixed points at infinity, one attracting and one repelling. We have already constructed fixed points with nonnegative and nonpositive strengths. Since the strength is nonzero (its absolute value is the stable translation length), then these fixed points correspond to what we want.

lemma *loxodromic-attracting-fixed-point*:
  assumes *loxodromic-isometry f*
  shows *attracting-fixed-point f* $\in$ *Gromov-boundary*
    Gromov-extension f (*attracting-fixed-point f*) = *attracting-fixed-point f*
    additive-strength f (*attracting-fixed-point f*) < *0*
$\langle proof \rangle$

lemma *loxodromic-repelling-fixed-point*:
  assumes *loxodromic-isometry f*
  shows *repelling-fixed-point f* $\in$ *Gromov-boundary*
    Gromov-extension f (*repelling-fixed-point f*) = *repelling-fixed-point f*
    additive-strength f (*repelling-fixed-point f*) > *0*
$\langle proof \rangle$

The attracting and repelling fixed points of a loxodromic isometry are distinct – precisely since one is attracting and the other is repelling.

lemma *attracting-fixed-point-neq-repelling-fixed-point*:
  assumes *loxodromic-isometry f*
  shows *attracting-fixed-point f* $\neq$ *repelling-fixed-point f*
$\langle proof \rangle$

The attracting fixed point of a loxodromic isometry is indeed attracting. Moreover, the convergence is uniform away from the repelling fixed point. This is expressed in the following proposition, where neighborhoods of the repelling and attracting fixed points are given by the property that the Gromov product with the fixed point is large.

The proof goes as follows. First, the Busemann function with respect to the fixed points at infinity evolves like the strength. Therefore, $f^n e$ tends to the repulsive fixed point in negative time, and to the attracting one in positive time. Consider now a general point x with $(\xi^-, x)_e \leq K$. This means that the geodesics from e to x and ξ^- diverge before time K. For large n, since $f^{-n}e$ is close to ξ^-, we also get the inequality $(f^{-n}e, x)_e \leq K$. Applying f^n and using the invariance of the Gromov product under isometries yields $(e, f^n x)_{f^n e} \leq K$. But this Gromov product is equal to $d(e, f^n e) - (f^n e, f^n x)_e$ (this is a general property of Gromov products). In particular, $(f^n e, f^n x) \geq d(e, f^n e) - K$, and moreover $d(e, f^n e)$ is large. Since $f^n e$ is close to ξ^+, it follows that $f^n x$ is also close to ξ^+, as desired.

The real proof requires some more care as everything should be done in ereal, and moreover every inequality is only true up to some multiple of δ. But everything works in the way just described above.

proposition *loxodromic-attracting-fixed-point-attracts-uniformly*:
assumes *loxodromic-isometry f*
shows $\exists N. \forall n \geq N. \forall x.$ *extended-Gromov-product-at basepoint x* (*repelling-fixed-point f*) $\leq$ *ereal K*
$\longrightarrow$ *extended-Gromov-product-at basepoint* (((*Gromov-extension f*)^^*n*) *x*) (*attracting-fixed-point f*) $\geq$ *ereal M*
⟨*proof*⟩

We deduce pointwise convergence from the previous result.

lemma *loxodromic-attracting-fixed-point-attracts*:
assumes *loxodromic-isometry f*
xi $\neq$ *repelling-fixed-point f*
shows ($\lambda n.$ ((*Gromov-extension f*)^^*n*) *xi*) $\longrightarrow$ *attracting-fixed-point f*
⟨*proof*⟩

Finally, we show that a loxodromic isometry has exactly two fixed points, its attracting and repelling fixed points defined above. Indeed, we already know that these points are fixed. It remains to see that there is no other fixed point. But a fixed point which is not the repelling one is both stationary and attracted to the attracting fixed point by the previous lemma, hence it has to coincide with the attracting fixed point.

theorem *loxodromic-unique-fixed-points*:
assumes *loxodromic-isometry f*
shows *Gromov-extension f xi* = *xi* $\longleftrightarrow$ *xi* = *attracting-fixed-point f* $\vee$ *xi* = *repelling-fixed-point f*
⟨*proof*⟩

end

References

[BH99] Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999. MR MR1744486

[BS00] Mario Bonk and Oded Schramm, *Embeddings of Gromov hyperbolic spaces*, Geom. Funct. Anal. **10** (2000), 266–306. MR MR1771428

[CDP90] Michel Coornaert, Thomas Delzant, and Athanase Papadopoulos, *Parties quasi-convexes d'un espace hyperbolique*, pp. 106–123, Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.

[GdlH90] Étienne Ghys and Pierre de la Harpe (eds.), *Sur les groupes hyperboliques d'après Mikhael Gromov*, Progress in Mathematics, vol. 83, Birkhäuser Boston Inc., Boston, MA, 1990, Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988. MR MR1086648

[Min05] Igor Mineyev, *Flows and joins of metric spaces*, Geom. Topol. **9** (2005), 403–482. MR 2140987

[Shc13] Vladimir Shchur, *A quantitative version of the Morse lemma and quasi-isometries fixing the ideal boundary*, J. Funct. Anal. **264** (2013), no. 3, 815–836. MR 3003738