Topology, Geometry, and Physics

John Morgan

University of Haifa, Israel

March 28 – 30, 2017

The Basics of Manifold Topology

Poincaré

While there were antecedents in the work of Guass, Riemann, Betti, Cauchy, and others, Poincaré's work from 1892 through 1905 in a series of 7 articles, *Analysis Situs* and its complements, established Topology as an independent sub-discipline within mathematics.

Figure: Poincaré

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

(iii) to find algebraic invariants of manifolds,

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

(iii) to find algebraic invariants of manifolds,

(iv) study the properties of those invariants,

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

(iii) to find algebraic invariants of manifolds,

(iv) study the properties of those invariants,

(v) to use the invariants to distinguish manifolds that are topologically distinct, and

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

(iii) to find algebraic invariants of manifolds,

(iv) study the properties of those invariants,

(v) to use the invariants to distinguish manifolds that are topologically distinct, and

(vi) eventually to use the invariants to classify all manifolds.

Poincaré's Approach

(i) To define the notion of higher dimensional manifolds,

(ii) to give various ways of presenting these manifolds,

(iii) to find algebraic invariants of manifolds,

(iv) study the properties of those invariants,

(v) to use the invariants to distinguish manifolds that are topologically distinct, and

(vi) eventually to use the invariants to classify all manifolds.

We will see how this program continues to inspire work in topology today and also how naïve Poincaré's vision was.

The Plan of the Lectures

My goal in these lectures is to show you some of the (fairly) recent developments in low dimensional topology, i.e., the topology of manifolds of dimensions 3 and 4 and how both geometry and physics influence our understanding of these manifolds. But to set the stage and to 'warm up' I will begin with a review of the classical, and well-kinown, theory of surfaces. After that I will review the topological classification of simply connected 4-manifolds and Donsaldson's smooth invariants. Next, we will discuss the Jones polynomial and Khovanov homology of knots in 3-space and approaches to these invariants using ideas from physics. We will finish with a discussion of 3-dimensional maniifolds – culminating with Perelman's proof of the Geometrization for 3-manifolds.

PART I. TOPOLOGY OF SURFACES

Definition of Surfaces

Poincaré certainly had in mind the case of surfaces, well-understood by the time he did his foundational work in topology.

Definition of Surfaces

Poincaré certainly had in mind the case of surfaces, well-understood by the time he did his foundational work in topology.

A surface is a (Hausdorff) topological space Σ with the property that every point $x \in \Sigma$ has a neighborhood U homeomorphic to the open unit ball B^2 in the plane.

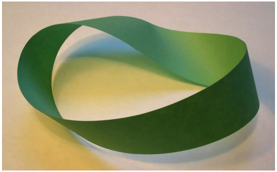
Definition of Surfaces

Poincaré certainly had in mind the case of surfaces, well-understood by the time he did his foundational work in topology.

A surface is a (Hausdorff) topological space Σ with the property that every point $x \in \Sigma$ has a neighborhood U homeomorphic to the open unit ball B^2 in the plane.

Given a homeomorphism $\varphi\colon U \to B^2$ we pull back the usual coordinates (x, y) on B^2 to functions, still called x and y, on U. These are local coordinates on Σ defined near x, and U together with its local coordinates is called coordinate patch. So a surface is a (Hausdorff) topological space with the property that it can be covered by local coordinate patches.

Examples of surfaces



Local Coordinates

We can cover Σ with local coordinate patches $\{U_a\}_{a\in A}$ with local coordiantes (x_a, y_a). On the overlap $U_a \cap U_b$, the functions x_b and y_b are continuous functions of (x_a, y_a), meaning that the overlap transition is a homeomorphism from an open subset of U_a to an open subset of U_b but in general nothing more can be said.

Local Coordinates

We can cover Σ with local coordinate patches $\{U_a\}_{a \in A}$ with local coordiantes (x_a, y_a). On the overlap $U_a \cap U_b$, the functions x_b and y_b are continuous functions of (x_a, y_a), meaning that the overlap transition is a homeomorphism from an open subset of U_a to an open subset of U_b but in general nothing more can be said.

One can impose more structure on the surface by requiring that we can cover the surface by a set of coordinate patches so that the overlaps are restricted in various ways. Examples are C^k-structures, C^∞-structures, real analytic structures, complex analytic structures, algebraic structures, and many others others.

Morse Functions on (smooth) Surfaces

It turns out that, for surfaces, it is no restriction to suppose that the surface in question has a C^∞-structure (called a smooth structure). We shall now make this assumption, which allows us to use calculus on surfaces.

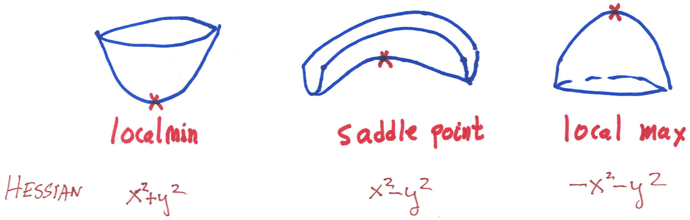
Morse Functions on (smooth) Surfaces

It turns out that, for surfaces, it is no restriction to suppose that the surface in question has a C^∞-structure (called a smooth structure). We shall now make this assumption, which allows us to use calculus on surfaces.

Let us consider a compact (smooth) surface. For simplicity let us suppose that it is smoothly embedded in a Euclidean space. Then choosing a generic direction to be 'height', the height function will have only isolated critical points and at each the Hessian of second derivatives of the function will be non degenerate:

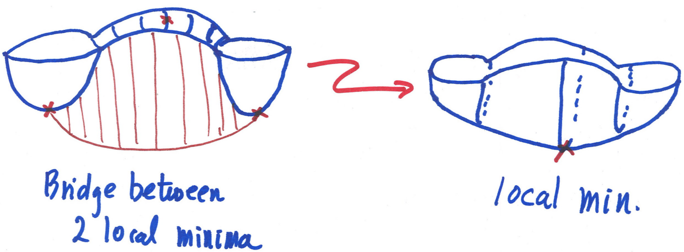
Morse Functions on (smooth) Surfaces

There are 3 possible types of critical points:



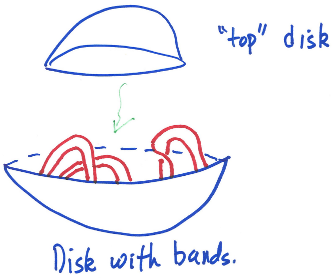
Combinatorial Picture of a surface

Restricting attention to connected surfaces, if there is more than one local minimum, then there is a bridge between two of them passing over a single critical point of index 1.



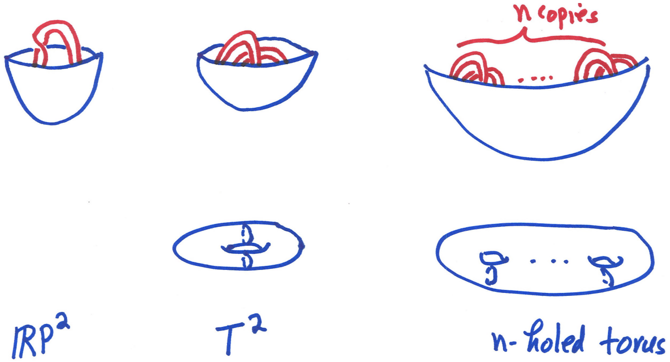
We can then 'push' the bridge down canceling the critical point one index one against one of the local minima.

Combinatorial Picture of a surface



Inductively, we can assume the height function has only one local minimum, and dually, only one local maximum. This leads to a picture of the surface as a disk with a certain number of 'bands' attached so that the boundary of the resulting surface is a single circle. Then a second disk is attached along that surface forming the compact surface.

Examples of Surfaces



Homology and Cohomology

Like any topological space, a surface has homology groups and cohomology groups. A Morse function can be used to produce a chain complex that computes these groups. One of the main properties of the homology and cohomology of a surface, indeed of any compact manifold, is that they satisfy Poincaré duality. In terms of a Morse function this duality is realized by turning the function over; i.e. replacing it by its negative. This sends a critical point of index k to a critical point of index $n - k$.

Interestingly the classification of compact surfaces agrees with the classification of finite dimension $\mathbb{Z}/2\mathbb{Z}$-vector spaces V with non-degenerate symmetric pairings to $\mathbb{Z}/2\mathbb{Z}$. The identification associates to a surface Σ the vector space $H^1(\Sigma;\mathbb{Z}/2\mathbb{Z})$ and the pairing is

$$a\otimes b\mapsto\langle a\cup b,[\Sigma]\rangle,$$

or equivalently to two homology classes it associates their homological intersection. Such a pairing is isomorphic to a diagonal pairing with 1s down the diagonal or to a direct sum of

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

The corresponding surfaces are a connected sum or real projective planes or a multi-holed torus. In the latter case the number of holes is called the genus.

PAT II. RIEMANNIAN GEOMETRY OF SURFACES

Riemannian metrics on surfaces

A Riemannian metric on a surface is a smoothly varying family of positive inner products on the tangent spaces. In local coordinates (x^1, x^2) we express the metric as

$$g_{ij}(x^1, x^2)dx^i \otimes dx^j,$$

where $g_{ij}(x^1, x^2)$ is a symmetric matrix of smooth functions that is positive definite at each point. Indeed,

$$g_{ij} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$$

is the inner product of the coordinate partial derivatives.

Riemannian metrics on surfaces

A Riemannian metric on a surface is a smoothly varying family of positive inner products on the tangent spaces. In local coordinates (x^1, x^2) we express the metric as

$$g_{ij}(x^1, x^2)dx^i \otimes dx^j,$$

where $g_{ij}(x^1, x^2)$ is a symmetric matrix of smooth functions that is positive definite at each point. Indeed,

$$g_{ij} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$$

is the inner product of the coordinate partial derivatives.

Every smooth surface, indeed every smooth manifold, has a Riemannian metric. Simply use a partition of unity to piece together standard Euclidean metrics on coordinate patches. Clearly, a surface has lots of Riemannian metrics, in fact an infinite dimensional space of them.

2^{nd} order approximation to a surface in 3-space: Curvature

Consider a surface Σ in 3-space. We can restrict the ambient Euclidean metric to define a Riemannian metric on Σ. Let $p \in \Sigma$. Translate and rotate the Euclidean coordinates of the ambient space so that locally near $p \in \Sigma$ the surface is given as the graph of a function $z = f(x, y)$ with p being the point $(0, 0, f(0, 0))$ and with $\nabla f(0, 0) = 0$. Then the tangent plane to Σ at p is the plane $\{z = 0\}$ and to second order the surface is given by

$$z(x,y) = f(0,0) + (x,y) \begin{pmatrix} \partial_{xx} f(0,0) & \partial_{xy} f(0,0) \\ \partial_{yx} f(0,0) & \partial_{yy} f(0,0) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

2^{nd} order approximation to a surface: Curvature

Rotating the x, y-coordinates allows us to assume that the matrix of partial derivatives of f at $(0,0)$ is diagonal – the new coordinate directions are called the directions of principle curvature, and the values of $-\partial_{xx} f(0,0)$ and $-\partial_{yy} f(0,0)$ are called the principle curvatures at p.

The product of the principle curvatures is called the Gauss curvature and is denoted K. It is of course the determinant of the matrix of second partials of f at $(0,0)$:

$$K = \det \begin{pmatrix} \partial_{xx} f(0,0) & \partial_{xy} f(0,0) \\ \partial_{xy} f(0,0) & \partial_{yy} f(0,0) \end{pmatrix}.$$

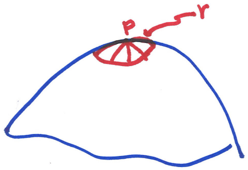
Gauss Curvature

The principle curvatures depend on the way the surface sits in 3-space but the Gauss curvature only depends on the Riemannian metric on the surface induced by the embedding in space, not the embedding itself.

In fact, we have

$$K(p) = \lim_{r \mapsto 0} \frac{\pi r^2 - \mathrm{Area}(B(p,r))}{\pi r^4/12},$$

where B(p,r) is the metric ball centered at p of radius r.



Gauss Curvature

That is to say the Gauss curvature measures the area defect (positive curvature) or area excess (negative curvature) of small balls centered at the point compared to the ball of the same radius in the plane.

Gauss Curvature

That is to say the Gauss curvature measures the area defect (positive curvature) or area excess (negative curvature) of small balls centered at the point compared to the ball of the same radius in the plane.

Indeed the area formula for the Gauss curvature of a surface in 3-space tells us how to generalize to any surface with a Riemannian metric – use the same area formula to define the Gauss curvature.

$$K(p) = \lim_{r \mapsto 0} \frac{\pi r^2 - \mathrm{Area}(B(p,r))}{\pi r^4/12}.$$

Gauss-Bonnet Theorem

There is a beautiful connection between the curvature and the topology of a surface:

Theorem

(Gauss-Bonnet Theorem) Let Σ be a compact surface and g a Riemannian metric on Σ with K_g its curvature. Then

$$\int_\Sigma K_g dvol = 2\pi\chi(\Sigma),$$

where $\chi(\Sigma)$ is the Euler characteristic of Σ.

Recall that

$$\chi(\Sigma) = \mathrm{rk}\, H^0(\Sigma) - \mathrm{rk}\, H^1(\Sigma) + \mathrm{rk}\, H^2(\Sigma).$$

Complex structures on surfaces

Let Σ be an oriented surface. A Riemannian metric determines a positive definite inner product on the tangent space at every point and hence an identification of the tangent space at every point with $\mathbb{C}$, up to rotation. $[SO(2) = U(1)]$. This determines a decomposition of the complexification of the cotangent space

$$T^*\Sigma \otimes_{\mathbb{R}} \mathbb{C} = T^{1,0}\Sigma \oplus T^{0,1}\Sigma,$$

where $T^{1,0}$ is the space of complex linear maps and $T^{0,1}$ is the space of complex anti-linear maps.

Complex structures on surfaces

This determines a decomposition of the differential d, which maps complex-valued functions on Σ to complex valued one-forms, as $d = \partial + \bar{\partial}$. It is a theorem that $\bar{\partial}$ determines a complex structure on Σ. Namely, near every point p there is a function z to the complex numbers with $\bar{\partial} z = 0$ and with $\partial z(p) \neq 0$. Such local functions determine local complex coordinates and make Σ a complex curve.

Universal Covering of a surface

Having imposed a complex structure on Σ let us consider the universal covering $\widetilde{\Sigma}$. It is a simply connected complex surface and has a Riemannian metric invariant under all complex automorphisms. Up to a constant rescaling, there are only three possibilities:

- S^2: the round metric
- $\mathbb{C}$: the Euclidean metric $dx^2 + dy^2$,
- the upper half-plane $\mathbb{H}$: the Poincaré metric $\frac{dx^2+dy^2}{y^2}$.

There is another model for the 3^{rd} example, namely the interior of the unit disk with the metric

$$\frac{4(dx^2 + dy^2)}{(1 - r^2)^2}.$$

This is also called the Poincaré metric.

Universal Covering of a surface

Since the group of complex automorphisms acts by isometries and acts transitively, it follows that these metrics are of constant curvature: 1, 0, and -1 respectively. Consequently,

Theorem

Any compact Riemann surface admits a metric of constant curvature $-1, 0$, or 1. If the surface is compact, the curvature of this constant curvature metric has the same sign as the Euler characteristic and the volume of the surface is 2π times the absolute value of the Euler characteristic.

Three types of surfaces

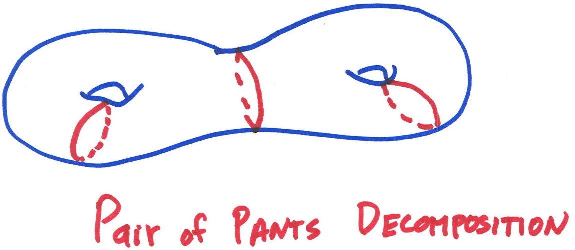
Round: S^2 and $\mathbb{R}P^2$.
Flat: T^2 and the Klein bottle
Negative or hyperbolic: all orientable surfaces of $g > 1$ and all connected sums of at least 3 projective planes.

Space of Flat Metrics on T^2: The Modular curve

Any complex structure on the torus is the quotient of $\mathbb{C}$ by a lattice. For the moment, fix a basis for the lattice. Modulo scaling and rotations we can assume that the lattice is generated by $\{1, \tau\}$ for some $\tau \in \mathbb{H}$. Changing the basis of the lattice produces an action of $SL(2, \mathbb{Z})$ (by linear fractional transformations) and the space of tori is identified with $\mathbb{H}/SL(2, \mathbb{Z})$. This is an interesting and much studied complex space, but we will not say more about it.

Space of hyperbolic metrics on a surface of genus $g > 1$

Let Σ be an orientable Riemann surface of genus $g > 1$ with a hyperbolic metric. Fix a system of $n = 3g - 3$ disjointly embedded loops $\{A_1, \ldots, A_n\}$ that divide the surface up into pairs of pants. We can make the A_i geodesic loops. Then we have the Fenchel-Nielsen coordinates for this metric: $\ell_1, \ldots, \ell_n$ are the lengths of the geodesics homotopic to $A_1, \ldots, A_n$ and $r_1, \ldots, r_n$ are rotation parameters along these loops.



Space of hyperbolic metrics on a surface of genus $g > 1$

This identifies the space of marked hyperbolic surfaces of genus g with $\mathbb{R}^{6g-6}$. Again the group of homotopy classes of surface automorphisms (called the mapping class group) acts on this space with finite stabilizers and the quotient is the moduli space of hyperbolic surfaces of genus g, or equivalently complex curves of genus g, another much studied space.

PART III: TOPOLOGY OF 4-MANIFOLDS

No Classification is Possible

There can be no classification of compact 4-manifolds of the Poincaré envisioned. The reason is that every finitely presented group occurs as the fundamental group of a compact 4-manifold, and it is a classical result that finitely presented groups cannot be classified.

For this reason, and for reasons of keeping life as simple as possible, we concentrate on simply connected 4-manifolds

Smooth versus topological

4 is the first dimension where there is a difference between smooth manifolds and topological manifolds. On the topological side there is a very good classification, at least for simply connected 4-manifolds.

Smooth versus topological

4 is the first dimension where there is a difference between smooth manifolds and topological manifolds. On the topological side there is a very good classification, at least for simply connected 4-manifolds.

Let us begin with the homotopy classification. The only homological invariant of such a manifold M is $H^2(M;\mathbb{Z})$, which is a free abelian group. Choosing an orientation on M determines a symmetric pairing

$$H^2(M;\mathbb{Z}) \otimes H^2(M;\mathbb{Z}) \to \mathbb{Z},$$

given by

$$a \otimes b \mapsto \langle a \cup b, [M] \rangle,$$

or if you prefer one can consider the homological intersection on dual group $H_2(M;\mathbb{Z})$. The isomorphism class of this pairing determines M up to homotopy equivalence.

Poincaré duality tells us that these pairings are unimodular. Thus, we can find a basis for $H^2(M;\mathbb{Z}) \otimes \mathbb{R} = H^2(M;\mathbb{R})$ in which the matrix for the intersection form is diagonal. We denote by $b_2^{\pm}(M)$ the number of positive and negative entries on the diagonal. The index of M is the signature of this pairing, i.e., $b_2^+(M) - b_2^-(M)$.

Freedman's classification

Theorem

(Freedman) Every symmetric, unimodular pairing occurs as the pairing of a compact, simply connected topological 4-manifold. If the pairing is even then the realizing simply connected manifold is unique up to homeomorphism. If the pairing is odd, then there are exactly two homeomorphism classes of simply connected, topological manifolds realizing the pairing and one of them is stably smooth in the sense that its product with $\mathbb{R}$ has a smooth structure.

As a corollary we have the 4-dimensional version of the Poincaré Conjecture.

Corollary

(Freedman) A compact, simply connected 4-manifold with the homology of S^4 is homeomorphic to the 4-sphere.

Smooth 4-manifolds

In contrast to this result, Donaldson first proved:

Theorem

(Donaldson) A definite even form is not the intersection form of any simply connected smooth manifold.

He went on to show:

Theorem

(Donaldson) There are non-diffeomorphic compact 4-manifolds that are homeomorphic.

In fact, using the same techniques one can show:

Theorem

There are infinitely many pairwise non-diffeomorphic 4-manifolds all of which are homeomorphic,

Principal G-bundles

All of these smooth theorems rely on understanding properties of the moduli space of solutions to the Anti-Self Dual equations for connections on principal $SU(2)$-bundles over the 4-manifold.
Recall for any group Lie group G, a principal G-bundle over a smooth manifold M is a smooth manifold P together with a smooth submersion $\pi\colon P \to M$ and a smooth free action $P \times G \to P$ with the property that π factors to give a smooth identification of the quotient space P/G with M.

Connections on Principal G-bundles

Definition

A connection on a principal G-bundle $P \to M$ is a G-invariant, smoothly varying family of ‘horizontal’ subspaces $H_p \subset P_p$.

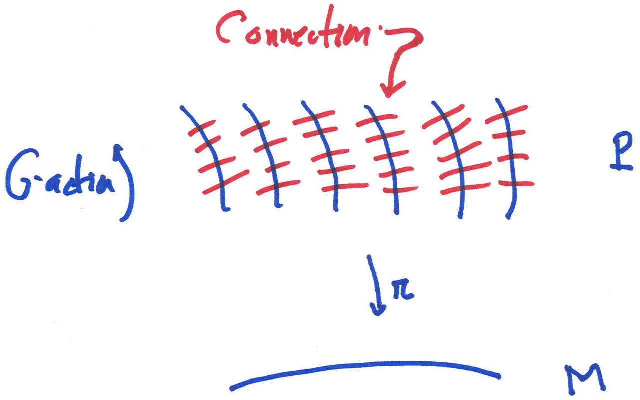
Connections on Principal G-bundles

Definition

A connection on a principal G-bundle $P \to M$ is a G-invariant, smoothly varying family of 'horizontal' subspaces $H_p \subset P_p$.

Horizontal means complementary to the tangent space to the fiber, or equivalently, mapping via $d\pi$ isomorphically onto $TM_{\pi(p)}$.

Connections on Principal G-bundles



Connections on Principal G-bundles

There are two ways to view a connection. One is parallel translation.

Connections on Principal G-bundles

There are two ways to view a connection. One is parallel translation.

A connection A allows us to define parallel translation along paths in the base. Suppose that $\gamma\colon [0,1] \to M$ is a smooth path from x to y. We define parallel translation

$$P_\gamma\colon \pi^{-1}(x) \to \pi^{-1}(y)$$

as follows. For any $p \in \pi^{-1}(x)$ there is a unique path $\widetilde{\gamma}_p$ that

(i) projects onto γ,

(ii) begins at p, and

(iii) has horizontal tangent vector at each point.

We define $P_\gamma(p) = \widetilde{\gamma}_p(1)$. This is a G-equivariant diffeomorphism from $\pi^{-1}(x)$ to $\pi^{-1}(y)$.
N.B. In general, parallel translation from $\pi^{-1}(x)$ to $\pi^{-1}(y)$ depends on the path γ connecting x and y.

Covariant Derivative

Parallel translation in the principal bundle determines parallel translation in any associated vector bundle $\mathcal{V} = P \times_G V$, where V is a (finite dimensional) linear representation of G. Namely, a curve in the total space of $\mathcal{V}$ is parallel if it is of the form $[\gamma(t), v]$ for a parallel path $\gamma(t)$ in P and a fixed vector $v \in V$. Parallel translation in a vector bundle allows us to define the covariant derivative

$$\nabla \colon \Omega^0(M; \mathcal{V}) \to \Omega^1(M; \mathcal{V})$$

as follows. Given a local section σ of $\mathcal{V}$ defined near $x \in M$ and given a tangent vector $\tau \in TM_x$ we express σ as $[\tilde{p}, \tilde{v}]$ where $\tilde{p}$ is a local section of $P \to M$ horizontal in the τ-direction and $\tilde{v}$ is a local function $M \to V$, and then we define

$$\nabla(\sigma)(\tau) = [\tilde{p}(x), \frac{\partial \tilde{v}(x)}{\partial \tau}].$$

ad

The Connection 1-form

The other way to view a connection is as a one-form on the principal bundle.

The Connection 1-form

The other way to view a connection is as a one-form on the principal bundle.
A connection A allows us to define a linear map $\omega_A\colon TP \to T_fP$, where T_fP means the subbundle tangent to the fibers of the projection to M. Furthermore, using the G action, we can identify T_fP with the Lie algebra $\mathfrak{g}$ of G. Thus, we have $\omega_A\colon TP \to \mathfrak{g}$. The G-invariance of the connection translates into an equivariance equation:

$$\omega_A(\tau g) = g^{-1}\omega_A(\tau)g.$$

The form ω_A does not descend to a one-form on M because it is non-trivial along the fibers. But since any two connection forms agree in the vertical direction, their difference vanishes on the fibers and satisfies the equivariance equation above. This means that the difference of two connections is a one-form on the base, M, with values in $\mathrm{ad}P$.

The Curvature of a Connection

We define the curvature of the connection as

$$F_A = d\omega_A + \frac{1}{2}[\omega_A, \omega_A].$$

This is a 2-form on P satisfying the equivariance property above. The Jacobi identity for G implies that this 2-form descends to a two-form on M with values in $\mathrm{ad}(P)$. It is the curvature 2-form.

The Curvature of a Connection

The curvature is the obstruction to the vanishing of ∇^2 in the following sense. As we have defined it

$$\nabla \colon \Omega^0(M; \mathrm{ad}P) \to \Omega^1(M; \mathrm{ad}P).$$

This extends by the Leibnitz rule to

$$\nabla \colon \Omega^i(M; \mathrm{ad}P) \to \Omega^{i+1}(M; \mathrm{ad}P),$$

by

$$\nabla(\omega \otimes \sigma) = d\omega \otimes \sigma + (-1)^{\mathrm{deg}(\omega)}\omega \wedge \nabla(\sigma).$$

Then

$$\nabla^2 \colon \Omega^0(M; \mathrm{ad}P) \to \Omega^2(M; \mathrm{ad}P)$$

turns out to be linear over the functions, and hence is multiplication by a 2-form with values in $\mathrm{ad}P$. That two form is the curvature 2-form.

Integrability

The vanishing of ∇^2 means that for any pair of vector fields X and Y near $p \in M$, we have

$$\nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X - \nabla_{[X,Y]} = 0.$$

This is exactly the integrability connection on the horizontal distribution. The connection is integrable if and only if its curvature vanishes, if and only if locally there is a trivialization of the principal bundle $P \to M$ in which the connection is the one induced by the product structure.

Integrability

The vanishing of ∇^2 means that for any pair of vector fields X and Y near $p \in M$, we have

$$\nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X - \nabla_{[X,Y]} = 0.$$

This is exactly the integrability connection on the horizontal distribution. The connection is integrable if and only if its curvature vanishes, if and only if locally there is a trivialization of the principal bundle $P \to M$ in which the connection is the one induced by the product structure.

For integral connections since $\nabla^2 = 0$ we get a complex of differential forms $\Omega^*(M; \mathrm{ad}P)$ with differential ∇ and we can then define the cohomology $H^*(M; \mathrm{ad}P)$, analogous to the deRham cohomology (which is the case when P is the trivial bundle with structure group $\mathbb{R}^*$.)

The ASD equations

Let M be a compact, connected, oriented Riemannian 4-manifold, and let $P \to M$ be a principal $SU(2)$-bundle. We consider the Yang-Mills energy of the connection given by

$$\frac{1}{4\pi^2}\int_M |F_A|^2 dvol.$$

In dimension 4, the Hodge $*$ operator on 2-forms squares to the identity and hence its eigenspaces determine a decomposition $\Lambda^2 T^*M$ as $\Lambda^+(M) \oplus \Lambda^-(M)$ and hence a decomposition of 2-forms as self-dual plus anti-self dual. These subspaces are orthogonal under the L^2-inner product. Hence,

$$\int_M |F_A|^2 dvol = \int_M (|F_A^+|^2 + |F_A^-|^2) dvol.$$

On the other hand, the Chern class $c_2(P)$ of the bundle is given by

$$\frac{1}{8\pi^2}\int_M \mathrm{tr}(F_A \wedge F_A) = \frac{1}{4\pi^2}\int_M (|F_A^-|^2 - |F_A^+|^2)dvol.$$

[The normalized positive definite inner product on $su(2)$ is $A \otimes B \mapsto -2\mathrm{tr}(AB)$.] If $\int_M c_2(P)$ is positive then the absolute minima of the energy function occurs when $F_A^+ = 0$, i.e., when the curvature is anti-self dual.

The moduli space of ASD connections

We suppose that $c_2(P) > 0$. Then the absolute minima are the ASD connections, namely $\{A \big| F_A^+ = 0\}$. The space of ASD connections on P is acted on by the automorphisms of the bundle P. For a generic metric on M the quotient space, the space of gauge equivalence classes of ASD connections which is denoted $\mathcal{M}(P)$, is smooth away from reducible connections and

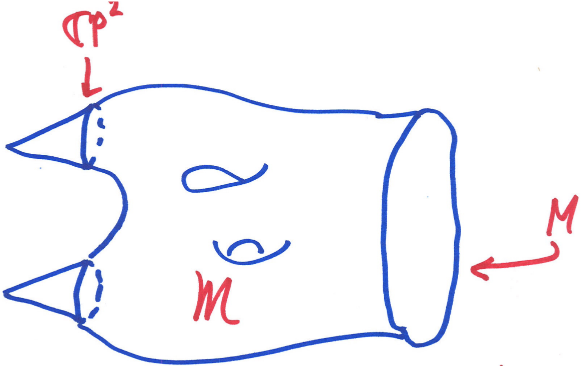
$$\dim \mathcal{M}(P) = 8\Big(\int_M c_2(P)\Big) - 3(1 + b_2^+(M)).$$

If $b_2^+(M) > 0$ then (for a generic metric) there are no reducible connections. For $b_2^+(M) > 1$, for a generic path of metrics there are no reducible connections.

The case $b_2^+ = 0$ and $c_2(P) = 1$

Consider now the case when M is a simply connected manifold with $b_2^+(M) = 0$ and $c_2(P) = 1$. Then the moduli space $\mathcal{M}(P)$ is 5 dimensional. Gauge equivalence classes of reducible connections are in natural one-to-one correspondence with the set of pairs of cohomology classes $\{\pm x\} \in H^2(M;\mathbb{Z})$ with $x^2 = -1$. Each reducible connection is a singular point of the moduli space whose neighborhood is homeomorphic to the cone of $\mathbb{C}P^2$. The moduli space is non-compact and a neighborhood of infinity in $\mathcal{M}(P)$ is diffeomorphic to $M \times [0, \infty)$. The ASD connections in this neighborhood are almost flat over almost all of M and have a 'bubble' of charge 1 concentrated near a point.

The moduli space of ASD connections



MODULI SPACE FOR $c_2 = 1$; $b_2^+(M) = 0$

Removing small neighborhoods around each reducible connection and adding a copy of M at infinity extracts from $\mathcal{M}(P)$ a compact oriented 5 manifold whose boundary is M together with one copy of $\mathbb{C}P^2$ for each pair $\{\pm x\}$ with $x^2 = -1$. It follows that the index of M must be the sum of integers ± 1, one for each pair $\{\pm x\}$ with $x^2 = -1$. This can happen only if the form is diagonalizable over the integers with -1s down the diagonal.

Corollary

(Donaldson) If a positive definite unimodular form is the intersection form of a compact, smooth, simply connected 4-manifold then the form is diagonalizable over the integers. In particular, no even positive definite form is the intersection form of such a manifold.

The results about non-uniqueness of smooth structures on certain topological 4-manifolds are proved in a similar way. One considers moduli spaces $\mathcal{M}(P)$ of gauge equivalence classes of ASD connections of bundles P of higher Chern class. These moduli spaces are compactified to $\overline{\mathcal{M}}(P)$ by adding idealized connections at infinity which record the limiting 'background connection' and how bubbling takes place. There is a natural map from $H_2(M) \to H^2(\overline{\mathcal{M}})$ and hence a map from the polynomial algebra generated by $H_2(M)$ to the cohomology ring of $\overline{\mathcal{M}}$. Integrating over the fundamental class of $\overline{\mathcal{M}}$ produces a homogeneous polynomial function on $H_2(M)$ whose degree is one-half the dimension of the moduli space, and hence which depends on the Chern class of P. These are the Donaldson polynomial invariants, which are well-defined independent of the metric if $b_2^+(M) > 1$. These invariants are used to distinguish non-diffeomorphic manifolds that are homeomorphic.

One computation

A $K3$ surface is a smooth quartic hypersurface in $\mathbb{C}P^3$. The Donaldson polynomials of the $K3$ surface are given by

$$D_{2n} = \frac{Q^n}{2^n},$$

where Q is the quadratic intersection form on H_2.

Equivalently, the Donaldson power series is $D = \exp(Q/2)$.

One computation

A $K3$ surface is a smooth quartic hypersurface in $\mathbb{C}P^3$. The Donaldson polynomials of the $K3$ surface are given by

$$D_{2n} = \frac{Q^n}{2^n},$$

where Q is the quadratic intersection form on H_2.

Equivalently, the Donaldson power series is $D = \exp(Q/2)$.

Producing a new, homeomorphic surfaces by doing log transforms along one or two fibers produces a surface with Donaldson invariants which are polynomials in Q and multiplies of the exceptional fibers, which then are different from the Donaldson polynomial of the $K3$. These give examples of homeomorphic, non-diffeomorphic surfaces.

The ASD equations are equations much studied by the physicists, and, after their use in this way in mathematics, physics produced a surprising twist. The information in the Donaldson polynomial invariants can also be obtained from a simpler set of equations called the Seiberg-Witten equations. These are equations where the gauge group is abelian and (at least in all known examples) the moduli spaces are zero dimensional. While the invariants carry equivalent information, technically the SW invariants have proved easier to deal with and are now routinely used instead of the Donaldson invariants.

The ASD equations are equations much studied by the physicists, and, after their use in this way in mathematics, physics produced a surprising twist. The information in the Donaldson polynomial invariants can also be obtained from a simpler set of equations called the Seiberg-Witten equations. These are equations where the gauge group is abelian and (at least in all known examples) the moduli spaces are zero dimensional. While the invariants carry equivalent information, technically the SW invariants have proved easier to deal with and are now routinely used instead of the Donaldson invariants.

The argument that SW equations give the same information to the Donaldson polynomials was originally a non-rigorous physics one using properties of quantum field theories which have no mathematical formulation to say nothing of mathematical proof. Now there are mathematically rigorous arguments covering many cases, but still no complete mathematical proof exists for this statement.

Situation for compact, simply connected smooth 4-manifolds

We have two sets of invariants of these manifolds: the cohomology H^2 its intersection pairing and the Donaldson invariants, or the (believed to be equivalent) Seiberg-Witten invariants. (There are also other invariants inspired by these but they are now known to carry the same information). The first set of invariants is equivalent to the homotopy type and cannot distinguish homeomorphic smooth manifolds. The second has had much success for certain classes of manifolds, for example algebraic surfaces and symplectic manifolds where they can be computed from the geometric structure.

Situation for compact, simply connected smooth 4-manifolds

We have two sets of invariants of these manifolds: the cohomology H^2 its intersection pairing and the Donaldson invariants, or the (believed to be equivalent) Seiberg-Witten invariants. (There are also other invariants inspired by these but they are now known to carry the same information). The first set of invariants is equivalent to the homotopy type and cannot distinguish homeomorphic smooth manifolds. The second has had much success for certain classes of manifolds, for example algebraic surfaces and symplectic manifolds where they can be computed from the geometric structure.

Still, we hardly know anything. We do not know whether the smooth version of the Poincaré Conjecture is true for 4-manifolds since the gauge theory invariants do not say anything about these manifolds.

Situation for compact, simply connected smooth 4-manifolds

We have two sets of invariants of these manifolds: the cohomology H^2 its intersection pairing and the Donaldson invariants, or the (believed to be equivalent) Seiberg-Witten invariants. (There are also other invariants inspired by these but they are now known to carry the same information). The first set of invariants is equivalent to the homotopy type and cannot distinguish homeomorphic smooth manifolds. The second has had much success for certain classes of manifolds, for example algebraic surfaces and symplectic manifolds where they can be computed from the geometric structure.

Still, we hardly know anything. We do not know whether the smooth version of the Poincaré Conjecture is true for 4-manifolds since the gauge theory invariants do not say anything about these manifolds.

We also do not even have a guess for a classification of compact, simply connected 4-manifolds.

Exotic 4-manifolds

There is a construction due to Fintushel-Stern which constructs for each knot in S^3 a smooth 4-manifold homeomorphic to the $K3$ surface (a quartic hypersurface in $\mathbb{C}P^3$).

Exotic 4-manifolds

There is a construction due to Fintushel-Stern which constructs for each knot in S^3 a smooth 4-manifold homeomorphic to the $K3$ surface (a quartic hypersurface in $\mathbb{C}P^3$).

The $K3$ surface is fibered by tori (with finitely many singular fibers) over S^2. Fintushel-Stern remove a neighborhood of a generic fiber of the form $T^2 \times D^2$ and glue in the product of a knot complement in S^3 times a circle in such a way that the boundary of a Seifert surface for the knot is glued to $\{pt\} \times \partial D^2$. They show that the result is homeomorphic to the $K3$ surface and the Seiberg-Witten invariants of the resulting manifold contain the Alexander polynomial of the knot and hence this invariant of the knot is captured by the smooth 4-manifold.

Exotic 4-manifolds

There is a construction due to Fintushel-Stern which constructs for each knot in S^3 a smooth 4-manifold homeomorphic to the $K3$ surface (a quartic hypersurface in $\mathbb{C}P^3$).

The $K3$ surface is fibered by tori (with finitely many singular fibers) over S^2. Fintushel-Stern remove a neighborhood of a generic fiber of the form $T^2 \times D^2$ and glue in the product of a knot complement in S^3 times a circle in such a way that the boundary of a Seifert surface for the knot is glued to $\{pt\} \times \partial D^2$. They show that the result is homeomorphic to the $K3$ surface and the Seiberg-Witten invariants of the resulting manifold contain the Alexander polynomial of the knot and hence this invariant of the knot is captured by the smooth 4-manifold.

It is not at all unreasonable to conjecture that distinct knots produce non-diffeomorphic 4-manifolds. If this is even close to being true, then one begins to sense the complexity of smooth 4-manifold theory and how little we understand.

PART IV: KNOT INVARIANTS

Consider a knot in 3-space or equivalently in the 3-sphere. A nice way to present a knot is by taking a planar projection and then indicating at each crossing which strand passes over and which passes under.

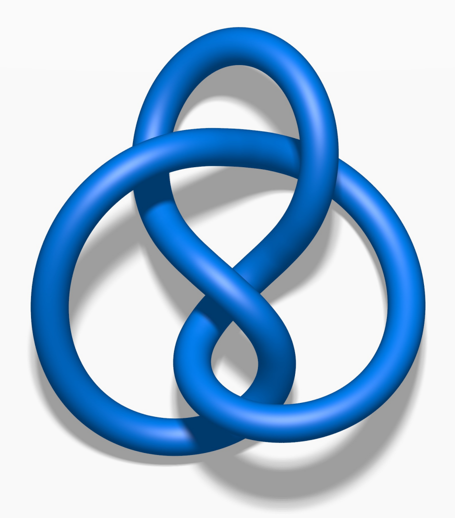


Figure: Figure Eight Knot

The Alexander Polynomial of a Knot

There is a classical invariant of the knot, called the Alexander polynomial, define by J. W. Alexander in 1927. Let $K \subset S^3$ be a knot and let $X = S^3 \setminus K$. We have $H_1(X; \mathbb{Z}) = \mathbb{Z}$, so that X has a unique infinite cyclic covering $\widetilde{X} \to X$. The homology $H_1(\widetilde{X}; \mathbb{Z})$ is a module over the ring $\Lambda = Z[t, t^{-1}]$ with the action of t being the map induced on H_1 by the generating deck transformation of $\widetilde{X}$. One shows that this module is cyclic and in fact can be written $\Lambda/(\Delta)$, where $\Delta \in \mathbb{Z}[t, t^{-1}]$ is defined up to multiplication by $\pm t^k$. The polynomial Δ is the Alexander polynomial.

There is a Skein relation defining the Alexander polynomial:

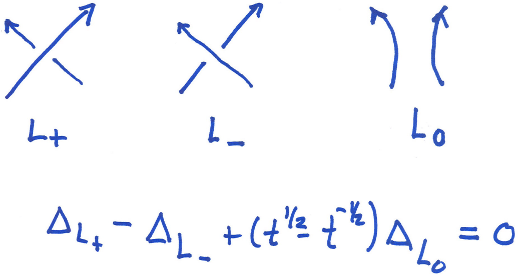


Figure: The Skein relation for the Alexander Polynomial

That together with the initialization that $\Delta(\text{trivial knot}) = 1$ determines the Alexander polynomial

The Jones Polynomial

Fifty years after the definition of the Alexander polynomial, Jones introduced a new polynomial invariant of knots. The Jones polynomial can be defined by the following Skein relation

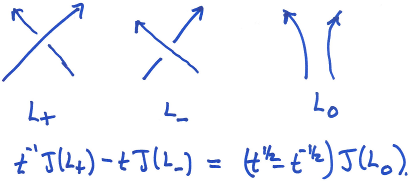


Figure: Skein relation for the Jones polynomial

That together with the initialization $J(\mathrm{trivial\ knot}) = 1$ determines the Jones polynomial.

The Jones Polynomial

Several things are not clear from the definition of the Jones polynomial. First of all, the definition uses a planar projection, and one has to show that it is an invariant of the knot, not the planar projection. There are a sequence of elementary moves connecting any one planar projection to any other, so one can prove that $J(t)$ is an invariant of the knot by showing that it is invariant under these moves. That is more or less what Jones did originally.

Also, this definition clearly only works for knots in the 3-sphere, again because of the use of a planar projection. It was not clear that this polynomial extends to an invariant of knots in more general 3-manifolds.

Khovanov homology

In 2000 Khovanov 'categorified' the Jones polynomial in the sense that he associated to a knot K in S^3 a bigraded chain complex $\oplus_{i,j} C^{i,j}$ with $d\colon C^{i,j} \to C^{i+1,j}$ whose homology $KH^{i,j}(K)$ is the Khovanov homology. The let $\chi^j = \sum_i (-1)^I \mathrm{rk} KH^{i,j}(K)$ be the Euler characteristic in the i-direction. Then we have $(q+q^{-1})^{-1} \sum_j \chi_j q^j$ is equal to $J(K)$, the Jones polynomial of the knot (with the substitution $q = t^{1/2}$).

Khovanov homology

In 2000 Khovanov ‘categorified’ the Jones polynomial in the sense that he associated to a knot K in S^3 a bigraded chain complex $\oplus_{i,j} C^{i,j}$ with $d\colon C^{i,j} \to C^{i+1,j}$ whose homology $KH^{i,j}(K)$ is the Khovanov homology. The let $\chi^j = \sum_i (-1)^I \mathrm{rk} KH^{i,j}(K)$ be the Euler characteristic in the i-direction. Then we have $(q+q^{-1})^{-1} \sum_j \chi_j q^j$ is equal to $J(K)$, the Jones polynomial of the knot (with the substitution $q = t^{1/2}$).
Khovanov homology is strictly stronger than the Jones polynomial in the sense that it distinguishes knots (and links) with the same Jones polynomial

The variable $q + q^{-1}$ in Jones' construction is replaced in Khovanov's by a graded free module A with a generators q, and q^{-1} of degree 1 and -1. Khovanov takes a planar projection of the knot, resolves all the crossings as in the skein relations and uses this to glue copies of tensor products of A together – basically doing a skein relation in the abelian category of modules over a given ring rather than in polynomials.

The variable $q + q^{-1}$ in Jones' construction is replaced in Khovanov's by a graded free module A with a generators q, and q^{-1} of degree 1 and -1. Khovanov takes a planar projection of the knot, resolves all the crossings as in the skein relations and uses this to glue copies of tensor products of A together – basically doing a skein relation in the abelian category of modules over a given ring rather than in polynomials.

The same issues that arise for the Jones polynomial arise here. The difficulty is in proving that the result is an invariant of the knot not the planar projection. Also, Khovanov homology is defined only for knots in S^3.

Physics approach to the Jones polynomial and Khovanov homology

One of the first connections between low dimensional topology and modern high energy theoretical physics was Witten's approach to the Jones polynomial from quantum field theory. He began with a topological quantum field theory based on the Chern-Simons functional. Associated to a connection A on the trivial G-bundle [take G compact and simple] over a compact 3-manifold we form we form

$$CS_k(A) = \frac{k}{4\pi} \int_M \mathrm{tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A).$$

Physics approach to the Jones polynomial and Khovanov homology

Unlike most gauge functions, e.g. the Yang-Mills functional, the Chern-Simons function does not depend on a metric on the manifold. It is purely topological. If we change the trivialization of the bundle by a map $M \to G$, then $CS_k(A)$ changes by $2\pi k$ times the degree of the map from $H_3(M) \to H_3(G)$ induced by the change of trivialization. This implies that as long as k is an integer, the action $\exp(iCS_k(A))$ is invariant.

Physics approach to the Jones polynomial and Khovanov homology

If $K \subset M$ is an oriented knot, then one adds to the action the trace of holonomy of the connection around K (trace in a fixed representation R of G). This is denoted $W_R(K) = \mathrm{tr}_W(\mathrm{hol}(A, K))$. The action with this 'operator' is then

$$W_R(K)\exp(i\int_M CS_k(A)).$$

As before this is a purely topological expression; there is no need to choose a metric or other auxiliary geometric data.

Physics approach to the Jones polynomial and Khovanov homology

Witten argues that this theory can be quatntized for any knot (or link) in any oriented 3-manifold (though one has to chose a framing on the tangent bundle of the 3-manifold). The case $G = SU(2)$ and R the two-dimensional representation and $M = S^3$ reproduces the Jones polynomial (or rather the values of this polynomial at roots of unity).

Physics approach to the Jones polynomial and Khovanov homology

Witten argues that this theory can be quatntized for any knot (or link) in any oriented 3-manifold (though one has to chose a framing on the tangent bundle of the 3-manifold). The case $G = SU(2)$ and R the two-dimensional representation and $M = S^3$ reproduces the Jones polynomial (or rather the values of this polynomial at roots of unity).

One advantage of Witten's approach is that it is manifestly 3-dimensional from the beginning – there is no choice of planar projection, and as a consequence he gets an extension to all 3-manifolds of a version of the Jones polynomial. One disadvantage is that this argument is not mathematically rigorous since it uses the full power of quantum field theory, but it has spurred mathematical developments.

Witten and Kapustin have written down geometric partial differential equations for a 5-dimensional theory whose moduli space they believe will produce the Khovanov homology. Much work is being done now by various mathematicians and physicists trying to show that these equations have the sort of properties that allow one to deal in a reasonable way with the moduli space of solutions. It is early days, but there is much interest and some real progress in turning these equations and their solutions into a useful mathematical theory.

PART V: TOPOLOGY OF 3-MANIFOLDS

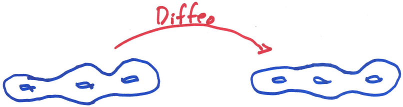
As is the case for surfaces, every 3-manifold has a smooth structure. For a Morse function F on M there are four types of critical points: those of index 0 (local minima), those of index 1, those of index 2, and those of index 3 (local maxima). As before we can arrange that there is a unique local min and a unique local max and that all the critical points of index 1 have smaller value of the function than all those of index 2. We then split the 3-manifold by the the level set $\Sigma = F^{-1}(t)$ for some value of t greater than the value of F at every critical point of index 1 and less than the value of F at every critical point of index 2.

Heegaard Decomposition

The surface $F^{-1}(t) = \Sigma$ splits M into $M_- \cup_\Sigma M_+$ where each of $M_\pm$ is obtained by adding solid handles to a 3-ball. For simplicity let us assume that M is orientable, Then Σ is orientable, i.e., Σ is a surface of genus $g \geq 0$ and each of $M_\pm$ is a solid handlebody with g handles.

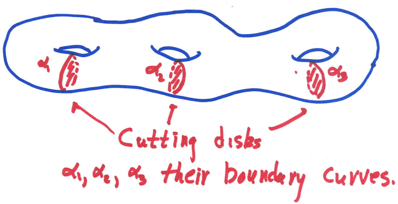
Heegaard Decomposition

The surface $F^{-1}(t) = \Sigma$ splits M into $M_- \cup_\Sigma M_+$ where each of $M_\pm$ is obtained by adding solid handles to a 3-ball. For simplicity let us assume that M is orientable, Then Σ is orientable, i.e., Σ is a surface of genus $g \geq 0$ and each of $M_\pm$ is a solid handlebody with g handles.



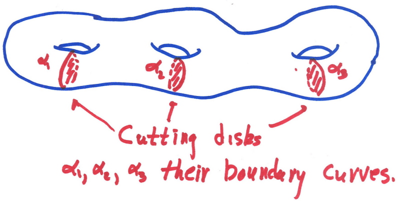
Heegaard Decomposition

The handlebody can be cut along g disks to produce a ball.



Heegaard Decomposition

The handlebody can be cut along g disks to produce a ball.



We can recover M from the surface Σ and two sets of g curves: $\{\alpha_1, \ldots, \alpha_g\}$ that bound disjoint cutting disks in M_- and $\beta_1, \ldots, \beta_g$ that bound the cutting disks in M_+. Each family is independently completely standard but together they contain the secret of the 3-manifold.

Heegaard Decomposition

This description of M is called a Heegaard decomposition. Its genus is the genus of the cutting surface. The problem is that a given 3-manifold has many different Heegaard decompositions. For example, M is topologically equivalent to S^3 if and only if it has a Heegaard decomposition of genus 0. But it has higher genus Heegaard decompositions which do not 'simplify' in any direct way to one of genus 0.

Topological Invariants of M^3

The only interesting homology group of a 3-manifold is its first homology group which is the abelianization of the fundamental group. Not surprisingly the fundamental group plays a central role in 3-manifold theory.

Topological Invariants of M^3

The only interesting homology group of a 3-manifold is its first homology group which is the abelianization of the fundamental group. Not surprisingly the fundamental group plays a central role in 3-manifold theory.

The fundamental group of M is the fundamental group of the splitting surface Σ modulo the normal subgroup generated by the curves $\alpha_1, \ldots, \beta_g$. Another way to think about this is to take the free group on $x_1, \ldots, x_g$ and then for each loop β_i form a word in the x_j by reading off, in order as one goes around β_i, the intersection points of β_i with the α_j (an intersection point with α_j adds the letter $x_j^{\pm 1}$ to the end of the word with the exponent recording the sign of the crossing).

The quotient of the free group generated by the x_i by the normal subgroup generated by the β-words is identified with the fundamental group of M.

Topological Invariants of M^3

The only interesting homology group of a 3-manifold is its first homology group which is the abelianization of the fundamental group. Not surprisingly the fundamental group plays a central role in 3-manifold theory.

The fundamental group of M is the fundamental group of the splitting surface Σ modulo the normal subgroup generated by the curves $\alpha_1, \ldots, \beta_g$. Another way to think about this is to take the free group on $x_1, \ldots, x_g$ and then for each loop β_i form a word in the x_j by reading off, in order as one goes around β_i, the intersection points of β_i with the α_j (an intersection point with α_j adds the letter $x_j^{\pm 1}$ to the end of the word with the exponent recording the sign of the crossing).

The quotient of the free group generated by the x_i by the normal subgroup generated by the β-words is identified with the fundamental group of M.

The first homology of the 3-manifold is the abelianization of this which can be read off from the matrix of homological intersections of the α and β curves in the first homology of Σ.

Conjecture

(Poincaré Conjecture) Let M be a compact 3-manifold, If M is simply connected (i.e., its fundamental group is trivial), then M is homeomorphic to the 3-sphere.

N.B. The converse is obvious.

Poincaré's suggested method of proof was to simplify the Heegaard decomposition, using the hypothesis of simple connectivity, to put the α and β curves in good position with respect to each other.

To date no one has been able to make that argument work, despite repeated attempts by many generations of topologists.

PART VI. Locally homogeneous Riemannian 3-manifolds

Riemannian curvature in higher dimensions

The analogue of Gauss curvature for higher dimensional manifolds is the Riemann curvature tensor. In each two-plane direction at each point there is a Gauss curvature. These fit together to produce a tensor with 4 indices $R_{ijk\ell}$ that is skew symmetric in i, j, skew symmetric in k, ℓ and symmetric in the interchange of i, j with k, ℓ. Thus, we can view R as a quadratic form on $\Lambda^2(TM)$. Given a pair of orthogonal unit vectors $\{e_1, e_2\}$ the value of the quadratic form on the element $e_1 \wedge e_2 \in \Lambda^2(TM)$ is called the sectional curvature in the 2-pane direction spanned by e_1 and e_2.

Riemannian curvature in higher dimensions

The analogue of Gauss curvature for higher dimensional manifolds is the Riemann curvature tensor. In each two-plane direction at each point there is a Gauss curvature. These fit together to produce a tensor with 4 indices $R_{ijk\ell}$ that is skew symmetric in i, j, skew symmetric in k, ℓ and symmetric in the interchange of i, j with k, ℓ. Thus, we can view R as a quadratic form on $\Lambda^2(TM)$. Given a pair of orthogonal unit vectors $\{e_1, e_2\}$ the value of the quadratic form on the element $e_1 \wedge e_2 \in \Lambda^2(TM)$ is called the sectional curvature in the 2-pane direction spanned by e_1 and e_2.

A manifold has constant curvature if all the eigenvalues of the quadratic form are equal, or equivalently if all the sectional curvatures are the same.. In all dimensions manifolds of constant negative curvature are hyperbolic. They are described as the quotient of the unit ball in n-space with its Poincaré metric divided out by a discrete, torsion-free group.

Riemannian curvature in dimension 3

Viewed as a quadratic form on $\Lambda^2 TM$, the Riemann curvature tensor can be diagonalized in an orthonormal basis. But this basis in general does not consist of two-plane directions.

But in dimension 3 we have the duality between TM and $\Lambda^2 TM$ which means that every element of $\Lambda^2 TM$ is a multiple of a two-plane direction. This implies that at each point there is an orthonormal frame $\{e_1, e_2, e_3\}$ such that the Riemann curvature tensor is diagonal in the basis $\{e_2 \wedge e_3, e_3 \wedge e_1, e_1 \wedge e_2\}$.

Definition

A Riemannian manifold M is homogeneous if its isometry group acts transitively. In this case $M = G/H$ where G is a Lie group and H is a compact subgroup. The metric on M is induced from an H-invariant metric on the Lie algebra $\mathfrak{g}$ of G by left translation.
A Riemannian manifold is locally homogeneous if any two points p and q have neighborhoods U_p and U_q that are isometric.

Lemma

If M is a compact (or complete) locally homogeneous manifold then its universal covering is homogeneous. In particular, there is a simply connected Lie group G, a compact subgroup H, and a discrete group Γ (either co-comapct or of co-finite volume) meeting H only in the identity element such that $M = \Gamma\backslash G/H$ with the metric being induced by left G-translation from an H-invariant metric on the Lie algebra $\mathfrak{g}$.

The list of simply connected Lie groups G containing a compact subgoup H of codimension 3 leads to the following exhaustive list of homogeneous 3-manifolds

- S^3, $\mathbb{H}^3$, and $\mathbb{R}^3$. These are the homogeneous 3-manifolds of constant curvature.
- $S^2 \times \mathbb{R}$ and $\mathbb{H} \times \mathbb{R}$. These are the homogeneous 3-manifolds that are products of (non-zero) constant curvature surfaces and the line
- The universal cover of $PSL(2, \mathbb{R})$
- The Heisenberg (nilpotent) group of matrices

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

- The solvable group $\mathbb{R}^* \ltimes \mathbb{R}^2$ where $t \in \mathbb{R}^*$ acts linearly with eigenvlaues $t^{\pm 1}$ on $\mathbb{R}^2$.

Compact locally homogeneous 3-manifolds of the various types are:

- Round (e.g. Lens spaces); flat, meaning finitely covered by a flat 3-torus; hyperbolic 3-manifolds
- $S^2 \times S^1$ or a manifold double covered by $S^2 \times S^1$; a hyperbolic surfaces times S^1 or a manifold finitely covered by such.
- a non-trivial circle bundle over a surface of genus > 1, or a manifold finitely covered by such.
- a non-trivial circle bundle over the torus, or a manifold finitely covered by such.
- a 2-torus bundle over the circle with Anosov monodromy.

Notice that all the various types are easy to list except for hyperbolic manifolds. Hyperbolic manifolds are in natural one-to-one correspondence with conjugacy classes of discrete, torsion-free co-compact subgroups of $PSL(2, \mathbb{C})$, which is the isometry group of hyperbolic 3-space. But there is no classification of these subgroups of $PSL(2, \mathbb{C})$.

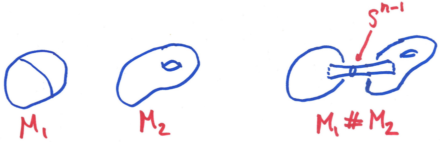
Thurston's Geometrization Conjecture

As a first guess, one might be tempted to say that every compact 3-manifold has a locally homogeneous metric. There is a simple reason why that is false. Except for $S^2 \times \mathbb{R}$, all the homogeneous 3-manifolds have trivial π_2. That means that if a locally homogeneous 3-manifold has non-trivial π_2 then it is either $S^2 \times S^1$ or double covered by this manifold.

Thurston's Geometrization Conjecture

As a first guess, one might be tempted to say that every compact 3-manifold has a locally homogeneous metric. There is a simple reason why that is false. Except for $S^2 \times \mathbb{R}$, all the homogeneous 3-manifolds have trivial π_2. That means that if a locally homogeneous 3-manifold has non-trivial π_2 then it is either $S^2 \times S^1$ or double covered by this manifold.

On the other hand, it is easy to construct lots of manifolds with non-trivial π_2 by taking connected sum. Given M_1 and M_2 we remove a ball from each and gluing the resulting manifolds with boundary together along their 2-sphere boundaries.



Connected sum

If M_1 nor M_2 are 3-manifolds, neither homotopy equivalent to S^3, e.g., if they each have non-trivial fundamental group, then the 2-sphere along which we glue the manifolds represents a non-trivial element in π_2 of the connected sum. Almost all such manifolds cannot be locally homogeneous. For example if M_1 and M_2 are non-simply connected and the order of $\pi_1(M_2)$ is at least 3, the result is not locally homogeneous.

Existence and Uniqueness of connected sum decomposition

It is a classical theorem in 3-manfiold topology that every 3-manifold decomposes as a connected sum of prime 3-manifolds, those that have no non-trivial connected sum decomposition. Furthermore, the prime factors are unique up to isomorphism. Thus, to classify all three manifolds it suffices to classify all prime 3-manifolds.

The Geometrization Conjecture

The Geometrrization Conjecture is now a theorem.

Theorem

Let M^3 be a compact, orientable prime 3-manifold. Then there exists a finite set of disjoint 2-tori and Klein bottles $\mathcal{T} = T_1, \ldots, T_k \subset M$, such that each component of $M \setminus \mathcal{T}$ has a locally homogeneous metric of finite volume. If we choose $\mathcal{T}$ to have a minimal number of connected components among all such collections, then the embedding of $\mathcal{T}$ into M is unique up to isotopy. The manifold M is determined by the topological type of components of $M \setminus \mathcal{T}$ and the isotopy classes of the gluings along the components T_i.

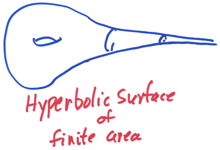
The Geomerization Conjecture includes as a special case the Poincaré Conjecture: If M is simply connected, the $\mathcal{T}$ must be empty and hence M has a locally homogeneous geometry. Since M is simply connected, this implies that it has a homogeneous geometry, which can only be the round metric on S^3.

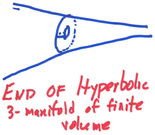
Ends of complete, locally homogeneous 3-manifolds of finite volume

A complete hyperbolic surface of finite volume has a finite number of ends. Each is diffeomorphic to $S^1 \times [0, \infty)$ and the cross-sectional length decreases exponentially as we go to the end.

Thus, any end of a circle bundle over a hyperbolic surface of finite area is diffeomorphic to $T^2 \times [0, \infty)$.
Similarly, each end of a hyperbolic 3-manifold of finite volume is diffeomorphic to $T^2 \times [0, \infty)$ and the diameter and area of the cross-sectional torus decrease at an exponential rate.

The Ends


Hyperbolic Surface
of
finite area


END OF Hyperbolic
3-manifold of finite
volume

PART VII: Perelman's proof of the Geometrization Conjecture

Ricci Flow

Perelman's approach is to use the Ricci flow equation introduced by Hamilton, who also established many of the essential properties of this flow. The Ricci flow equation is:

$$\frac{\partial g}{\partial t} = -2\mathrm{Ric}(g).$$

Let us understand the terms in this equation. We have a fixed manifold M and a smoothly varying family of Riemannian metrics $g(t)$. Recall that $g(t)$ is a symmetric contravariant 2-tensor; in local coordinates it is $g_{ij}dx^i \otimes dx^j$ with g_{ij} being a symmetric matrix of smooth functions of the coordinates $(x^1, \ldots, x^n)$. It has the additional property of being positive definite at every point.

Given the smooth family $g(t)$ of symmetric 2-forms we can differential with respect to t. The result is a symmetric contravariant 2-tensor $\Delta_{ij} dx^i \otimes dx^j$ with Δ_{ij} being a symmetric matrix of smooth functions

$$\Delta_{ij} = \frac{\partial g_{ij}(\mathbf{x}, t)}{\partial t}.$$

This symmetric 2-form is no longer necessarily positive definite.

Given the smooth family $g(t)$ of symmetric 2-forms we can differential with respect to t. The result is a symmetric contravariant 2-tensor $\Delta_{ij} dx^i \otimes dx^j$ with Δ_{ij} being a symmetric matrix of smooth functions

$$\Delta_{ij} = \frac{\partial g_{ij}(\mathbf{x}, t)}{\partial t}.$$

This symmetric 2-form is no longer necessarily positive definite.

What about the right-hand side of

$$\frac{\partial g}{\partial t} = -2\mathrm{Ric}(g)?$$

Ricci Curvature

The Ricci curvature is a symmetric contravariant 2-tensor, meaning in a local coordinate system it is of the form $\mathrm{Ric}_{ij} dx^i \otimes dx^j$ with Ric_{ij} a symmetric matrix of smooth funtions. It is obtained by tracing the Riemannian curvature tensor $R_{ijk\ell}$ on the middle two inidices j, k.

For a 3-manifold, as we have already observed, at each point p there is an orthonormal frame $\{e_1, e_2, e_3\}$ for the tangent space so that the Riemann curvature is diagonalized meaning the $\{e_1 \wedge e_2, e_2 \wedge e_3, e_3 \wedge e_1\}$ is a basis for $\Lambda^2 TM_p$ in which the Riemannian curvature tensor, viewed as a quadratic form on $\Lambda^2 TM_p$, is diagonal. Let $\lambda_3, \lambda_1, \lambda_2$ be the sectional curvatures on these three planes. Then the Ricci curvature is diagonal with respect to the basis $\{e_1, e_2, e_3\}$ of TM_p and is given by

$$\begin{pmatrix} \lambda_2 + \lambda_3 & 0 & 0 \\ 0 & \lambda_1 + \lambda_3 & 0 \\ 0 & 0 & \lambda_1 + \lambda_2 \end{pmatrix}$$

Short-time Existence and Uniqueness

According to Hamilton, given a compact Riemannian manifold (M, g_0) there is a $T > 0$ and a solution to the Ricci flow equation $(M, g(t))$ defined for $0 \leq t < T$ with $g(0) = g_0$. This solution is unique in the sense that given two solutions with the same initial condition they agree on their common interval of definition.

This result follows from general PDE theory and the fact that modulo the action of the diffeomorphism group the Ricci flow equation is a strictly parabolic equation.

Examples of Ricci flow

Consider the round metric g_0 on S^n with constant Ricci curvature 1 (i.e., constant sectional curvature $1/(n-1)$). Then $g(t) = (1-2t)g_0$, and the flow becomes singular at $t = 1/2$ when the metric shrinks to zero.

Consider a hyperbolic metric g_0 on M with constant Ricci curvature $-1/(n-1)$. Then $g(t) = (1+2t)g_0$ and the flow exists for all positive time and the hyperbolic manifold inflates by a constant (for each t) factor $\sqrt{1+2t}$.

Perelman's approach

For any compact 3-manifold M, the first step is to choose (arbitrarily) a Riemannian metric g_0 on M. Then apply Ricci flow with this initial condition. In the best circumstances, this flow will exist for all time $0 \leq t < \infty$ and as $t \mapsto \infty$ the manifold will decompose into pieces whose geometry and/or topology we can understand.

Perelman's approach

For any compact 3-manifold M, the first step is to choose (arbitrarily) a Riemannian metric g_0 on M. Then apply Ricci flow with this initial condition. In the best circumstances, this flow will exist for all time $0 \leq t < \infty$ and as $t \mapsto \infty$ the manifold will decompose into pieces whose geometry and/or topology we can understand.

There are two issues to confront:

(i) In general, the flow will not exist for all time but will have finite-time singularities – we must understand how to extend the flow past these to define a 'Ricci flow with surgeries' for all time, and

(ii) make the analysis as $t \mapsto \infty$.

The most difficult part of Perelman's analysis is understanding qualitatively the finite time singularities that develop in Ricci flow and extending the flow past them. Hamilton showed that given a Ricci flow $(M, g(t)),\ 0 \leq t < T_0 < \infty$ then the solution extends to an interval $[0, T_1)$ with $T_1 > T_0$ unless the curvature is unbounded as $t \mapsto T_0$. Perelman showed that f the curvature is unbounded as $t \mapsto T_0$ then there are three possible types of finite time singularities that can be occurring:

- Components that shrink to a point at some finite time, and as they shrink to a point the curvature approaches round. In particular, these components have a metric of constant positive curvature.
- Tubes diffeomorphic to $S^2 \times (-L, L)$ where the cross sectional 2-spheres have large positive curvature.
- Regions diffeommorphic the either B^3 or to $\mathbb{R}P^3 \setminus B^3$ of large positive curvature with ends $S^2 \times [0, L)$ as in the previous case.

Ricci flow with surgery

Let (M, g_0) be a compact Riemannian 3-manifold and let $(M, g(t)), 0 \leq t < T_0$ be the maximal Ricci flow with these initial conditions. It $T_0 < \infty$ there are singularities, as described in the previous slide forming as $t \mapsto T_0$. We define a manifold $M(T_0)$ as follows

- As we approach T_0, we remove any component shrinking to a point. These have round metrics so removing them does not affect whether the Geometrization Conjecture holds.
- For components of the second type we remove the center of the tube and cap off the two spheres by 3-balls of positive curvature. This does a connected sum decomposition.
- A component of the third type is either topologically a three-ball or $\mathbb{R}P^3 \setminus B^3$. We remove this component and cap off the end with a 3-ball. This either does not affect the topology or removes a connected sum with $\mathbb{R}P^3$.

Away from these regions we take the limiting metric. This defines $M(T_0), g(T_0))$

Ricci flow with surgery

It follows from the description on the previous slide, that if $M(T_0)$ satisfies the Geometrization Conjecture then the same is true of the manifold before surgery. Now we restart the Ricci flow at time T_0 with $(M(T_0), g(T_0)$ as the initial conditions at time T_0. We continue this process inductively, producing a Ricci flow with singularities defined for all positive time $(M(t), g(t), 0 \leq t < \infty$, whose initial condition is the is (M, g_0). If we can show that the manifold $M(t)$ for t sufficiently large satisfies the geometrization hypothesis then so does M.

It follows from the description on the previous slide, that if $M(T_0)$ satisfies the Geometrization Conjecture then the same is true of the manifold before surgery. Now we restart the Ricci flow at time T_0 with $(M(T_0), g(T_0))$ as the initial conditions at time T_0. We continue this process inductively, producing a Ricci flow with singularities defined for all positive time $(M(t), g(t), 0 \leq t < \infty$, whose initial condition is the is (M, g_0). If we can show that the manifold $M(t)$ for t sufficiently large satisfies the geometrization hypothesis then so does M.

So we turn to the analysis of $M(t)$ for $t >> 1$.

Scales

Consider a 3-dimensional Ricci flow with surgery $(M(t), g(t))$. We denote points in this flow by (x, t) meaning that $x \in M(t)$. The ball in $M(t)$ centered at x with radius r (with respect to $g(t)$) is denoted $B(x, t, r)$.

Scales

Consider a 3-dimensional Ricci flow with surgery $(M(t), g(t))$. We denote points in this flow by (x, t) meaning that $x \in M(t)$. The ball in $M(t)$ centered at x with radius r (with respect to $g(t)$) is denoted $B(x, t, r)$.

Definition

We define the Euclidean volume constant of a ball $B(x, t, \rho)$ to be the constant w with the property that $\mathrm{vol}(B(x, t, \rho) = w\rho^3$. This is invariant under rescaling. We say that a ball is w-collapsed if its Euclidean volume constant is less than w.

Scales

Consider a 3-dimensional Ricci flow with surgery $(M(t), g(t))$. We denote points in this flow by (x, t) meaning that $x \in M(t)$. The ball in $M(t)$ centered at x with radius r (with respect to $g(t)$) is denoted $B(x, t, r)$.

Definition

We define the Euclidean volume constant of a ball $B(x, t, \rho)$ to be the constant w with the property that $\mathrm{vol}(B(x, t, \rho) = w\rho^3$. This is invariant under rescaling. We say that a ball is w-collapsed if its Euclidean volume constant is less than w.
We define the negative curvature scale $\rho(x, t)$ to be the supremum of the $\rho > 1$ such that all sectional curvatures on $B(x, t, \rho)$ are bounded below by $-\rho^{-2}$. This is also a scale-invariant notion. This simply means that if we rescale the ball to have radius 1, then all the sectional curvatures are bounded below by -1.

Limits as $t \mapsto \infty$

As the next two propositions indicate, the natural metric to use as $t \mapsto \infty$ is $(1/t)g(t)$

Proposition

Fix a constant $w > 0$. Then for any $r > 0$ and any $\epsilon > 0$ there is $T < \infty$ such that if $t > T$ and if $r\sqrt{t}$ is less than the negative curvature scale $\rho(x,t)$ and if $B(x,t,r\sqrt{t})$ has Euclidean volume constant at least w then for the rescaled metric $\frac{1}{t}g(t)$ the Ricci curvature is within ϵ of -1. Furthermore, given $A < \infty$ if t is greater than a constant $T(A)$ then the same is true for all (y,s) with $y \in B(x,t,Ar\sqrt{t})$ and $s \in [t, t + Ar^2 t]$.

Limits as $t \mapsto \infty$

As the next two propositions indicate, the natural metric to use as $t \mapsto \infty$ is $(1/t)g(t)$

Proposition

Fix a constant $w > 0$. Then for any $r > 0$ and any $\epsilon > 0$ there is $T < \infty$ such that if $t > T$ and if $r\sqrt{t}$ is less than the negative curvature scale $\rho(x,t)$ and if $B(x,t,r\sqrt{t})$ has Euclidean volume constant at least w then for the rescaled metric $\frac{1}{t}g(t)$ the Ricci curvature is within ϵ of -1. Furthermore, given $A < \infty$ if t is greater than a constant $T(A)$ then the same is true for all (y,s) with $y \in B(x,t,Ar\sqrt{t})$ and $s \in [t, t + Ar^2 t]$.

Proposition

For any $w > 0$ there is $\bar{\rho} = \bar{\rho}(w) > 0$ such that for all t sufficiently large if $\rho(x,t) < \bar{\rho}\sqrt{t}$ then the Euclidean volume of $B(x,t,\rho(x,t))$ is less than w.

Limits as $t \mapsto \infty$

Fix $w > 0$ and define $M_{-}(t, w)$ to be those (x, t) for which the Euclidean volume constant of $B(x, t, \rho(x, t)$ is less than w, and we set $M_{+}(t, w)$ equal to its complement in $M(t)$. By the second proposition, for any $(x, t) \in M_{+}(t, w)$ the negative curvature scale is at least $\bar{\rho}(w)\sqrt{t}$. Thus, for every (x, t) in $M_{+}(w, t)$ the first proposition applies to show that for t sufficiently large after rescaling the metric by t^{-1} the curvature on the ball $B(x, \bar{\rho}(w))$ is every where close to -1 and this statement remains true for a time interval of length $A\rho(x, t)$. It follows that $M_{+}(t, w)$ with the rescaled metric $t^{-1}g(t)$ converges to (possibly disconnected) hyperbolic manifold of finite volume.

Limits as $t \mapsto \infty$

Now let us consider $M_-(t, w)$. It is w-collapsed on its negative curvature scale. Rescaling to make the negative curvature scale 1, we have a ball of radius 1 with sectional curvature bounded below by -1 which is volume collapsed. Such balls are Gromov-Hausdorff close to balls in spaces of dimension 0, 1 or 2.

Limits as $t \mapsto \infty$

Now let us consider $M_-(t,w)$. It is w-collapsed on its negative curvature scale. Rescaling to make the negative curvature scale 1, we have a ball of radius 1 with sectional curvature bounded below by -1 which is volume collapsed. Such balls are Gromov-Hausdorff close to balls in spaces of dimension 0, 1 or 2.

Points at which the approximating space is zero dimensional can be rescaled again to converge to a flat manifold.

Limits as $t \mapsto \infty$

Now let us consider $M_-(t, w)$. It is w-collapsed on its negative curvature scale. Rescaling to make the negative curvature scale 1, we have a ball of radius 1 with sectional curvature bounded below by -1 which is volume collapsed. Such balls are Gromov-Hausdorff close to balls in spaces of dimension 0, 1 or 2.

Points at which the approximating space is zero dimensional can be rescaled again to converge to a flat manifold.

At points where the approximating space is 2-dimensional one shows that $M_-(t, w)$ is fibered by circles over a ball (or more generally Seifert fibered).

Limits as $t \mapsto \infty$

Now let us consider $M_-(t, w)$. It is w-collapsed on its negative curvature scale. Rescaling to make the negative curvature scale 1, we have a ball of radius 1 with sectional curvature bounded below by -1 which is volume collapsed. Such balls are Gromov-Hausdorff close to balls in spaces of dimension 0, 1 or 2.

Points at which the approximating space is zero dimensional can be rescaled again to converge to a flat manifold.

At points where the approximating space is 2-dimensional one shows that $M_-(t, w)$ is fibered by circles over a ball (or more generally Seifert fibered).

At points where the approximating space is 1 dimensional, one shows that $M_-(t, w)$ is a fibration over the interval or circle with fibers either S^2, or T^2.

The local description can be pieced together to give a decomposition of $M_{-}(t, w)$ for w sufficiently small. The decomposition is along tori and Klein bottles into pieces that are Seifert fibered, $T^2 \times I$, and 2-torus bundles over the circle, and compact flat manifolds. All these have locally homogeneous geometries of finite volume (not necessarily coming from the Ricci flow, but rather from the topological classification of such manifolds).

The local description can be pieced together to give a decomposition of $M_-(t, w)$ for w sufficiently small. The decomposition is along tori and Klein bottles into pieces that are Seifert fibered, $T^2 \times I$, and 2-torus bundles over the circle, and compact flat manifolds. All these have locally homogeneous geometries of finite volume (not necessarily coming from the Ricci flow, but rather from the topological classification of such manifolds).

This completes the (outline) of the proof of the Geometrization Conjecture.

THANK YOU