Ingemar Bengtsson

ANTI-DE SITTER SPACE

Fall 1998

Anti-de Sitter space is the maximally symmetric solution of Einstein's equations with an attractive cosmological constant included; in reality the cosmological constant is certainly not attractive, but it is possible to regard it merely as a kind of regularisation of the long-distance behaviour of gravity. The conformal boundary of asymptotically anti-de Sitter space differs dramatically from that of asymptotically flat spacetimes, and this feature is usually crucial whenever anti-de Sitter space appears in mathematical physics. Notable examples are Friedrich's treatment of isolated systems in GR, the BTZ black holes, and also various studies in supersymmetry and string theory. (One can probably prove a theorem to the effect that string theory has an ergodic property that will make it come arbitrarily close to any point in idea space, if one waits long enough.)

This course was meant as a leisurely introduction to the geometry of anti-de Sitter space. The contents became:

- Quadric surfaces
- Hyperbolic spaces
- Anti-de Sitter space
- Asymptotia
- Green functions

I am sorry that the pictures were not computerized. Some of them were exercises, anyway.[1]

[1] Note added in 2009: Over the years, I have added a few pictures, but not enough. Those I did add are old ones.

QUADRIC SURFACES

Our spaces

Anti-de Sitter space belongs to a wide class of homogeneous spaces that can be defined as quadric surfaces in flat vector spaces. Since every quadratic form can be diagonalized over the reals, only diagonal quadrics need be considered. The signature is relevant however. Everybody knows that the n dimensional sphere $\mathbf{S}^n$ can be defined as the positive definite quadric

$$X_1^2 + \ ... \ + X_{n+1}^2 = R^2 \tag{1}$$

embedded in an Euclidean $n+1$ dimensional space. A sphere has constant positive curvature proportional to the inverse of R^2. We will almost always set the constant R and its analogue for other quadrics equal to one from now on, thus making the letter R available for other purposes. It is nevertheless worthwhile to remember that, unlike flat spaces, spaces with non-zero constant curvature have a natural length scale. As Gauss once wistfully remarked: "I have sometimes in jest expressed the wish that Euclidean geometry is not true. For then we would have an absolute a priori unit of measurement."

There are $n(n+1)/2$ Killing vectors in the embedding space that leave the sphere invariant, namely

$$J_{\alpha\beta} = X_\alpha \partial_\beta - X_\beta \partial_\alpha \ . \tag{2}$$

These are the Killing vectors of the sphere, and they generate the rotation group $SO(n+1)$. Given two points on the sphere it is always possible to go from one to the other along integral curves of some Killing vector field, which is why the sphere is said to be a homogeneous space.

Now suppose we change a sign in the quadric so that it becomes a hyperboloid rather than a sphere:

$$X_1^2 + \ ... \ + X_n^2 - U^2 = -1 \tag{3}$$

Because we also changed the sign on the right hand side this is a hyperboloid of two sheets. Just by looking at it it is clear that we have obtained a space

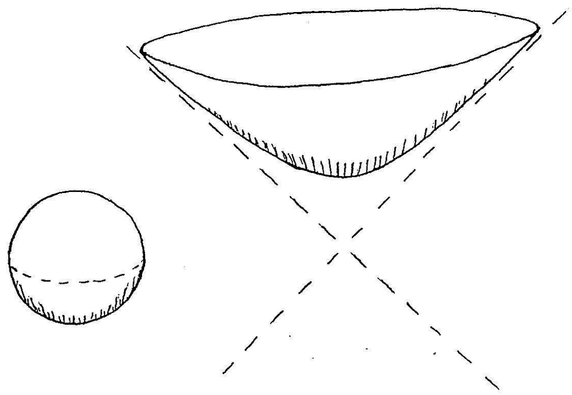

Figure 1: The sphere and the hyperbolic plane.

whose curvature is not constant. Although it is straightforward to write down $n(n+1)/2$ vector fields that leave it invariant, some of them fail to be Killing vectors with respect to the metric induced by the embedding. To be precise, this is the case for the n vector fields

$$J_{iU} = X_i\partial_U + U\partial_i \ , \tag{4}$$

where the index i runs from 1 to n. However, let us also change a sign in the metric on the embedding space so that it becomes a flat Minkowski space:

$$ds^2 = dX_1^2 + \ ... \ + dX_n^2 - dU^2 \ . \tag{5}$$

This metric is left invariant by all of the $n(n+1)/2$ transformations that leave the hyperboloid invariant, and it follows that their generators are Killing vectors of the induced metric on the hyperboloid as well. The group that they generate is $SO(n,1)$, which is the group of Lorentz transformations in the embedding space. Because of the way the quadric sits inside the Minkowski space the intrinsic metric induced on it has Euclidean signature, and because it is now a maximally symmetric space its curvature is necessarily constant and turns out to have a negative sign. We define hyperbolic space $\mathbf{H}^n$ as the upper sheet of such a hyperboloid.

If we consider a one sheeted hyperboloid

$$X_1^2 + \ ... \ + X_n^2 - X_{n+1}^2 = 1 \tag{6}$$

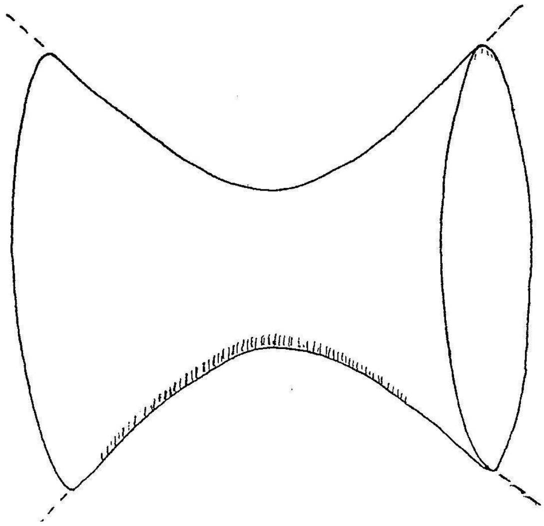

Figure 2: Anti-de Sitter space.

again embedded in Minkowski space we obtain a space with a Lorentzian metric of constant curvature. This is known as de Sitter space $\mathbf{dS}_n$.

To obtain anti-de Sitter space we change the rules yet again. Anti-de Sitter space $\mathbf{adS}_n$ is defined as the quadric

$$X_1^2 + \ ... \ + X_{n-1}^2 - U^2 - V^2 = -1 \tag{7}$$

embedded in a flat $n+1$ dimensional space with the metric

$$ds^2 = dX_1^2 + \ ... \ + dX_{n-1}^2 - dU^2 - dV^2 \ . \tag{8}$$

We can draw a picture of a two dimensional anti-de Sitter space, which is a hyperboloid of one sheet embedded in a three dimensional Minkowski space. It is (at least almost) evident from the picture that the intrinsic metric will have Lorentzian signature. The curve that goes around the waist of the hyperboloid is a closed timelike curve. This is not so important, since we can always "unwrap" the hyperboloid by going to the covering space. Note that in 1 + 1 dimensions we can always switch the meaning of timelike and spacelike. Then we obtain de Sitter space $\mathbf{dS_2}$, that has a closed space but no closed timelike curves. In general the topology of $\mathbf{adS}_n$ is $\mathbf{R}^{n-1} \otimes \mathbf{S}^1$ and the topology of $\mathbf{dS}_n$ is $\mathbf{S}^{n-1} \otimes \mathbf{R}$, so that it is only in two dimensions that de

Sitter space and anti-de Sitter space are thus simply related. Note also that, appearances notwithstanding, there is nothing special about the waist, since $\mathbf{adS}_n$ is in fact a homogeneous space having $n(n+1)/2$ Killing vectors that generate the group $SO(n-1,2)$. The waist looks special in the picture only because when looking at the picture one tends to interpret the hyperboloid as a surface in an Euclidean space, whereas in fact it is a surface in a Minkowski space.

Geodesics

As you may have noticed, the embedding coordinates are not very helpful for visualisation. Nor is the effort worthwhile since there are alternative intrinsic parametrisations that are much better in this regard. On the other hand the embedding coordinates do tend to simplify almost every calculation that one may want to do. As an example of this, consider geodesics. Let us think about anti-de Sitter space for definiteness, and use the notation

$$X^\alpha = (X, Y, Z, U, V) \qquad X_\alpha = (X, Y, Z, -U, -V) \tag{9}$$

$$X^2 = X^\alpha X_\alpha \qquad X_1 \cdot X_2 = X_1^\alpha X_{2\alpha} \ . \tag{10}$$

Using an overdot to denote a derivative we then take the Lagrangian

$$L = \frac{1}{2}\dot{X}^2 + \Lambda(X^2 + 1) \ . \tag{11}$$

The Lagrange multiplier term ensures that the geodesic is confined to the hyperboloid. We observe that there is a conserved tensor

$$k_{\alpha\beta} = X_\alpha \dot{X}_\beta - X_\beta \dot{X}_\alpha \ . \tag{12}$$

It obeys

$$k_{\alpha\beta}k^{\alpha\beta} = -2\dot{X}^2 \qquad k_{[\alpha\beta}k_{\gamma\delta]} = 0 \ . \tag{13}$$

It is easy to show conversely that given an anti-symmetric tensor obeying the latter condition there will be a geodesic which will be timelike, spacelike or lightlike depending on the value of $k_{\alpha\beta}k^{\alpha\beta}$. Further,

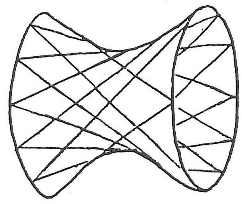

Figure 3: Null geodesics in $\mathbf{adS}_2$, courtesy Sir Christopher Wren.

$$k_{\alpha\beta}X^\beta = \dot{X}_\alpha \qquad k_{\alpha\beta}\dot{X}^\beta = \dot{X}^2 X_\alpha \qquad \Rightarrow \qquad \ddot{X}_\alpha = \dot{X}^2 X_\alpha \ . \tag{14}$$

Since $\dot{X}^2$ is a conserved quantity it is easy to solve this equation. In particular, for a lightlike geodesic it is as simple as it could be:

$$\ddot{X}^\alpha = 0 \ . \tag{15}$$

Hence lightlike geodesics appear as straight lines in the embedding space. The existence of two such straight lines through any given point in the hyperboloid of one sheet (such as $\mathbf{adS}_2$) is a surprising fact discovered by Sir Christopher Wren. The general solution for a spacelike geodesic is

$$X^\alpha = m^\alpha e^{\sqrt{\dot{X}^2}\tau} + n^\alpha e^{-\sqrt{\dot{X}^2}\tau} \ , \tag{16}$$

where m^α and n^α are constant vectors that obey

$$m^2 = n^2 = 0 \qquad\qquad 2m \cdot n = -1 \ , \tag{17}$$

and we have used the fact that $\dot{X}^2$ is constant along a geodesic. With minor changes this formula gives the general solution for a timelike geodesic as well.

We can use our formula to compute the geodesic distance d between any two points, that is to say the length of a geodesic connecting the points. For a spacelike geodesic (say) we find that

$$X(\tau_1) \cdot X(\tau_2) = m \cdot n (e^{\sqrt{\dot{X}^2}(\tau_1 - \tau_2)} + e^{\sqrt{\dot{X}^2}(\tau_2 - \tau_1)}) \ . \tag{18}$$

But $\sqrt{\dot{X}^2}(\tau_2 - \tau_1)$ is the integral of ds along a connecting geodesic, hence this product is equal to d. Therefore we find for the geodesic distance between two points with spacelike separation that (in an obvious notation)

$$\cosh d = -X_1 \cdot X_2 \ . \tag{19}$$

In a similar way we find for the geodesic distance that can be connected by a timelike geodesic that

$$\cos d = -X_1 \cdot X_2 \ . \tag{20}$$

Note that it can happen that $X_1 \cdot X_2 > 1$. If so, it is not possible to connect the pair of points with a geodesic.

The geodesic trajectories in $\mathbf{adS}_2$ can be visualized with very little effort. Let us switch the meaning of space and time and consider two dimensional de Sitter space instead, because both $\mathbf{dS}_2$ and $\mathbf{H}^2$ can be embedded in three dimensional Minkowski space, so that we can treat both cases simultaneously. We set

$$X^\alpha = (X, Y, U) \tag{21}$$

and observe that we now have a conserved vector

$$k_\alpha = \epsilon_{\alpha\beta\gamma} X^\beta \dot{X}^\gamma \tag{22}$$

which obeys

$$k \cdot X = 0 \ . \tag{23}$$

The trajectory of a geodesic in $\mathbf{dS}_2$ is therefore given by the intersection of the hyperboloid with a plane through the origin, and it will be a lightlike, timelike or spacelike geodesics depending on how the plane is situated. In particular it is now evident that all spacelike geodesics—all timelike geodesics in $\mathbf{adS}_2$—that start out from a given point will come together again in a point that lies halfway around the hyperboloid, and it is also evident that there will be points "on the other side" of the hyperboloid that cannot be connected with a given point by means of a geodesic. For the hyperbolic plane there is of course only one kind of geodesics, since we must demand that the vector k_α be spacelike to ensure that the plane meets the hyperboloid.

The two dimensional picture generalizes nicely. Let us first define totally geodesic surfaces (or submanifolds, if their dimension is larger than two): A totally geodesic surface is a surface such that any geodesic that is tangent to the surface at a point stays in the surface, or—which is the same thing—a surface such that any curve in the surface which is a geodesic with respect to the induced metric on the surface is also a geodesic in the space in which the surface is embedded. One can show that a surface (or submanifold) is totally geodesic if and only if its extrinsic curvature tensor—its second fundamental form—vanishes. In a flat spacetime it is obvious that any hyperplane is totally geodesic.

In a spacetime which is defined as a quadric in a flat embedding space any intersection of that quadric with a hyperplane through the orgin in the flat embedding space is totally geodesic, so we get a rich supply of such surfaces without effort. The point to remember is that in anti-de Sitter space any surface of the form

$$a \cdot X = 0 \tag{24}$$

is totally geodesic. Totally geodesic surfaces have the convenient property that the intersection of two such surfaces is itself totally geodesic. A one dimensional intersection of totally geodesic surfaces is a geodesic.

Projective quadrics

Now we have defined anti-de Sitter space and its cousins as quadric surfaces in flat vector spaces, and we have used this description to solve for the Killing vectors and the geodesics. But we have so far said nothing about their properties near infinity, where the embedding coordinates break down. This brings us to the conformal properties of these spaces, and the appropriate arena in which to understand these turns out to be a projective space. Let us stick to the two dimensional case for simplicity. There are then three spaces of constant curvature to be understood, namely $\mathbf{S}^2$, $\mathbf{E}^2$ and $\mathbf{H}^2$. The sphere can in fact be regarded as a null quadric in projective three-space, while the Euclidean and hyperbolic planes are subsets of the same quadric. To see this, recall that $\mathbf{RP}^3$ can be defined by means of homogeneous coordinates in $\mathbf{R}^4$; thus a point in $\mathbf{RP}^3$ is defined by

$$(X, Y, Z, U) \sim \lambda(X, Y, Z, U) \ , \qquad \lambda \neq 0 \ . \tag{25}$$

Here it is understood that quadruples of coordinates that differ by a common factor represent the same point in $\mathbf{RP}^3$. Geometrically the projective points are represented as straight lines through the origin in $\mathbf{R}^4$. Now consider the null cone

$$X^2 + Y^2 + Z^2 - U^2 = 0 \ . \tag{26}$$

Since this equation is homogeneous in the coordinates it is a well defined surface in $\mathbf{RP}^3$. Our claim is that this surface is related to the spaces that we wish to study.

In effect we are studying the space of generators of the light cone in Minkowski space. To understand its topology we take its intersection with the spacelike plane defined by

$$U = 1 \ . \tag{27}$$

It is clear that every generator of the light cone cuts this plane exactly one, and it is also clear that the intersection has the topology of a two-sphere, since it is given by the equation

$$X^2 + Y^2 + Z^2 = 1 \ . \tag{28}$$

The claim so far is that the space of generators of the light cone in Minkowski space has the topology of a sphere.

Alternatively we can intersect the light cone with a timelike plane, say

$$Z = 1 \ . \tag{29}$$

The intersection is now given by a two sheeted hyperboloid in a three dimensional space, namely

$$X^2 + Y^2 - U^2 = -1 \ . \tag{30}$$

Topologically this is two copies of the hyperbolic plane. But now we are missing all the generators that lie in the plane $Z = 0$, corresponding to an equator on the sphere. Finally we can intersect the light cone with a lightlike plane, such as

$$U = Z + 1 \ . \tag{31}$$

All the generators except the one given by $U = Z$ intersect this plane. The intersection of the plane with the quadric can be coordinatized by X and Y, which are unrestricted because

$$X^2 + Y^2 - (Z + U)(U - Z) = 0 \quad \Rightarrow \quad X^2 + Y^2 = Z + U \ . \tag{32}$$

Topologically this is a two dimensional plane. This time we are missing only one generator, so that in some sense the plane is a sphere with one point—the point "at infinity"—missing. The missing point is known as the conformal boundary of $\mathbf{E}^2$; similarly the conformal boundary of $\mathbf{H}^2$ is the missing circle of generators in the plane $Z = 0$. The conformal compactification of a given space includes its conformal boundary, which means that the conformal compactification of the Euclidean plane is a sphere while the conformal compactification of the hyperbolic plane is a closed disk. As we will see later on the discrepancy between the two cases is the key to understanding the difference between Euclidean and hyperbolic geometry.

Conformal properties

We have explained how $\mathbf{S}^2$, $\mathbf{H}^2$ and $\mathbf{E}^2$ are related as sets, but we have said nothing about metric properties. Now the null quadric is invariant under the group $SO(3,1)$ and therefore it singles out a natural metric in $\mathbf{R}^4$, namely the $SO(3,1)$ invariant Minkowski metric

$$ds^2 = dX^2 + dY^2 + dZ^2 - dU^2 \ . \tag{33}$$

The embeddings then induce metrics on the various intersections that we considered. The result is precisely the standard metrics on $\mathbf{S}^2$, $\mathbf{H}^2$ and $\mathbf{E}^2$. There is however some arbitrariness involved in this definition, since the selection of the various planes was arbitrary to some extent. Let us consider in more detail what kind of metrics that we can define on the space of generators in this manner. We introduce spherical polar coordinates

$$X = R\sin\theta\cos\phi \qquad Y = R\sin\theta\sin\phi \qquad Z = R\cos\theta \ . \tag{34}$$

The Minkowski metric becomes

$$ds^2 = dR^2 + R^2(d\theta^2 + \sin^2\theta d\phi^2) - dU^2 \ . \tag{35}$$

The equation that defines the lightcone takes the form

$$R^2 - U^2 = 0 \ . \tag{36}$$

Therefore the metric that is induced on the light cone by the embedding is

$$ds^2 = R^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{37}$$

This metric is degenerate on the three dimensional lightcone, which is at it should be since distances measured along the generators vanish. However, if the factor R^2 were fixed in some way we could choose to regard this metric as a metric on the two dimensional space of the generators, as parametrized by the angular coordinates. As such it is a non-degenerate metric and in fact it is a metric on a sphere. If we select one representative point on each generator through the condition $U = 1$ the factor R^2 becomes equal to unity and we obtain the standard "round" metric.

The point now is that there is nothing sacred about the condition $U = 1$. All that is needed to select one representative point on each generator is to intersect the light cone with some spacelike plane lying above (say) the origin. Therefore we can consider an entire family of conditions, each of which breaks the SO(1,3) invariance, given by

$$U = \Omega(\theta, \phi) \quad \Rightarrow \quad R = \Omega(\theta, \phi) \ . \tag{38}$$

If g_{ab} is the metric on the space of generators that is obtained through the choice $U = 1$, we get a new metric $\hat{g}_{ab}$ that obeys

$$\hat{g}_{ab} = \Omega^2 g_{ab} \ . \tag{39}$$

The factor in front is called a conformal factor, the metrics are said to be conformally related to each other, and changing the conformal factor of a metric is known as a conformal rescaling. An equivalence class of conformally related metrics is known as a conformal structure, and we can therefore conclude that it is only the conformal structure on the space of generators that is defined naturally by the metric on $\mathbf{R}^4$. Indeed the metrics on $\mathbf{S}^2$, $\mathbf{H}^2$ and $\mathbf{E}^2$ are related by conformal rescalings. (This is a triviality since it happens to be true for all two dimensional spaces; but we can make the same statement for $\mathbf{S}^3$, $\mathbf{H}^3$ and $\mathbf{E}^3$, and then it is non-trivial.)

Let us nevertheless settle for the standard round metric on $\mathbf{S}^2$, that is to say let us intersect the lightcone with the plane $U = 1$. The natural invariance group that acts in the problem is the group of Lorentz transformations in $\mathbf{R}^4$. A Lorentz transformation maps generators of the lightcone to other generators, and hence points of $\mathbf{S}^2$ to other points. If the Lorentz transformation is a boost it will also change the spacelike plane $U = 1$ to some other spacelike plane, and therefore the corresponding transformation of $\mathbf{S}^2$ must be accompanied by a conformal rescaling of the metric leaving its conformal structure unchanged. The group of transformations of a two dimensional space that preserve a given conformal structure is known as the conformal group $C(2)$, so that in effect we have proved that

$$C(2) = SO_0(3,1) \ . \tag{40}$$

The subscript denotes the connected component of the group. The terminology is a possible source of confusion here: A conformal transformation is a map from points to other points that is either an isometry or an isometry accompanied by a conformal rescaling—it is a very different concept from that of a conformal rescaling alone. A conformal transformation is a diffeomorphism, while a conformal rescaling is not.

An infinitesimal conformal transformation takes place along a vector field ξ^α that obeys the equation

$$\mathcal{L}_\xi g_{\alpha\beta} = \alpha g_{\alpha\beta} \ , \tag{41}$$

where $\mathcal{L}$ denotes the Lie derivative and α is some scalar function. The vector field is called a conformal Killing vector. The two-sphere admits six linear independent conformal Killing vectors that generate the Lorentz group $SO(3,1)$. It is easy to show that—locally—two conformally related spaces admit the same number of conformal Killing vectors, although a vector field that generates an isometry in the one case may be accompanied by a conformal rescaling in the other. Nevertheless the conformal group does not act on $\mathbf{E}^2$ or $\mathbf{H}^2$ because of global problems. In the case of the Euclidean plane this is easily remedied: From the point of view of the projective quadric it is evident that the problem is that there are Lorentz transformations that map a given generator to the "missing generator" that is parallell to the lightlike plane used to select $\mathbf{E}^2$. Conversely some Lorentz transformations map the missing generator into $\mathbf{E}^2$. Hence the conformal group does act on the conformal completion of $\mathbf{E}^2$, that is to say on $\mathbf{E}^2$ with the missing generator added as a point at infinity. In the case of $\mathbf{H}^2$ more elaborate measures would be required to achieve the same end.

There is nothing in the discussion so far that is special to two dimensions. Hence we can say immediately that the connected component of the conformal group $C(p,q)$ of a space of constant curvature with a metric of signature (p,q) is isomorphic to $SO_0(p+1,q+1)$. As a matter of fact there are some quite special features connected with conformal transformations in two dimensions; we will return to those. We will also—briefly—return to discuss the topology of the conformal completion of Minkowski space, which is a bit more elaborate than that of a sphere.

Exercises:

- A sphere is moving past you at relativistic speed. Show that *pace* length contraction it looks like a sphere.

HYPERBOLIC GEOMETRY

Advertisement

Hyperbolic geometry is of interest to us because anti-de Sitter space permits a natural foliation by means of hyperbolic spaces, because hyperbolic space is the natural analytic continuation of anti-de Sitter space, and because of the light it sheds on the importance of the conformal boundary of a space. Unlike anti-de Sitter space hyperbolic space is also of considerable direct physical interest: At the moment there is fairly strong evidence that the geometry of the universe (in the Friedmann approximation) is hyperbolic. This explains why we will go into considerable detail in this chapter.

Coordinates

We have defined hyperbolic space as the upper sheet of a two sheeted hyperboloid embedded in a Minkowski space. For visualisation a suitable choice of intrinsic coordinates is much better. It should be kept firmly in mind that coordinates do not matter, only the geometry itself does. The basic idea behind coordinate systems is to set up a one-to-one correspondence between the points of the space one is interested in and the points of some subset of $\mathbf{R}^n$. Quite literally, we are making a map of the space. Now one has a strong intuition for the properties of three dimensional Euclidean space $\mathbf{E}^3$, and one tends to interpret the subset of $\mathbf{R}^n$ that occurs in the coordinate map through one's intuition about Euclidean space. Therefore, if some particular aspect of the space one is interested in can be mimicked by the properties of some subset of three dimensional Euclidean space, the coordinates will also be helpful for intuitive understanding. However, if the space we are interested in is in fact not Euclidean space, there is always a danger that one will be led astray by coordinate based intuition. The best antidote to any tendency to think that coordinates are important is to change them often.

Stereographic projection

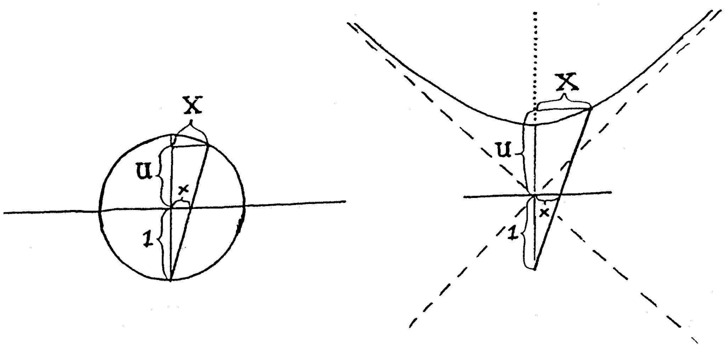

Figure 4: Stereographic projections of the sphere and the hyperbolic plane.

With this *caveat* we come to the coordinate system that I think is the best one for visualizing hyperbolic space, namely the stereographic coordinates. We can treat $\mathbf{S}^3$ and $\mathbf{H}^3$ simultaneously; they are defined as the quadrics

$$X^2 + Y^2 + Z^2 \pm U^2 = \pm 1 \tag{42}$$

where we let the upper sign stand for $\mathbf{S}^3$ and the lower for $\mathbf{H}^3$ throughout. The idea is now to perform a projection from the point $(X, Y, Z, U) = (0, 0, 0, -1)$ to the hyperplane at $U = 0$. In the case of the sphere this means that one point will be missing from the map, but this is unavoidable. In this way we obtain

$$x = \frac{X}{U+1} \qquad y = \frac{Y}{U+1} \qquad z = \frac{Z}{U+1} \; . \tag{43}$$

Also

$$\rho^2 \equiv x^2 + y^2 + z^2 = \frac{X^2 + Y^2 + Z^2}{(1+U)^2} = \frac{\pm 1 \mp U^2}{(1+U)^2} = \frac{\pm 1 \mp U}{1+U} \; , \tag{44}$$

where the upper sign applies to $\mathbf{S}^3$, the lower to $\mathbf{H}^3$. Equivalently the stereographic coordinates obey

$$X = \frac{2x}{1 \pm \rho^2} \tag{45}$$

$$Y = \frac{2y}{1 \pm \rho^2} \tag{46}$$

$$Z = \frac{2z}{1 \pm \rho^2} \tag{47}$$

$$U = \frac{1 \mp \rho^2}{1 \pm \rho^2} \ . \tag{48}$$

The intrinsic metric becomes

$$ds^2 = \frac{4}{(1 \pm \rho^2)^2}(dx^2 + dy^2 + dz^2) \ . \tag{49}$$

Let us concentrate on hyperbolic space, since the reader is supposed to know about the sphere already. Then the range of the coordinates is restricted by the condition $\rho < 1$, as is apparent from the fact that the conformal factor of the metric diverges at $\rho = 1$. So we have a picture of $\mathbf{H}^3$ as the interior of the unit ball in Euclidean space. This representation is known as the Poincaré ball, while its even more famous counterpart in two dimensions is known as the Poincaré disk. An important advantage of this coordinate system is that the metric is manifestly conformally flat, that is to say that it differs from the flat metric only through a conformal rescaling. An interesting consequence of this is that if we draw a picture using the coordinates as Cartesian coordinates on flat Euclidean space then all angles will be correctly given in the picture, since angles are not changed by conformal rescalings. Distances on the other hand are distorted by the projection; the distortion is smallest around the origin, so we can think of the stereographic coordinates as a kind of "magnifying glass" that gives an accurate picture of the region around the origin.

Other coordinate systems

Another very useful coordinate system represents $\mathbf{H}^3$ as a half space in coordinate space. Set

$$Y = \frac{y}{x} \qquad Z = \frac{z}{x} \tag{50}$$

$$X + U = \frac{1}{x} \tag{51}$$

$$X - U = -\frac{x^2 + y^2 + z^2}{x} \ , \tag{52}$$

where $x > 0$. This is the half space representation of $\mathbf{H}^3$ with the intrinsic metric taking the form

$$ds^2 = \frac{1}{x^2}(dx^2 + dy^2 + dz^2) \ . \tag{53}$$

Like the stereographic coordinates this coordinate system has the advantage that the metric is manifestly conformally flat. It is often superior in calculations.

Many other coordinate systems are useful in order to discuss various aspects of hyperbolic geometry. Like Felix Klein we may prefer to view $\mathbf{H}^3$ as a subset of $\mathbf{RP}^3$, or as the space of lines through the origin in $\mathbf{R}^4$. This suggests that we should adopt the origin as the point from which to project the hyperboloid onto some plane. If the plane of projection is taken to be $U = 1$ this results in

$$X = \frac{\rho' \cos\phi \sin\theta}{\sqrt{1 - \rho'^2}} \qquad Y = \frac{\rho' \sin\phi \sin\theta}{\sqrt{1 - \rho'^2}} \tag{54}$$

$$Z = \frac{\rho' \cos\theta}{\sqrt{1 - \rho'^2}} \qquad U = \frac{1}{\sqrt{1 - \rho'^2}} \ , \tag{55}$$

where the intrinsic coordinates are spherical polars. Hyperbolic space now appears as the Klein ball, with the metric

$$ds^2 = \frac{d\rho'^2}{(1 - \rho'^2)^2} + \frac{\rho'^2}{1 - \rho'^2}(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{56}$$

This is no longer manifestly conformally flat, so angles are distorted in the Klein ball. On the other hand—since totally geodesic surfaces are the intersections of the hyperboloid with planes through the origin in the embedding space—it is clear from the construction that totally geodesic surfaces now appear as flat planes in the ball, and geodesics as straight lines. This is an advantage if we want to impose boundary conditions for differential equations on totally geodesic surfaces and then solve them numerically, since it will be easy to make coordinate grids.

The Klein ball can be obtained from the Poincaré ball through a rescaling of the radial coordinate. Other rescalings of ρ are useful too, such as defining r through

$$dr = \frac{2d\rho}{1-\rho^2} \ . \tag{57}$$

Then the metric becomes (in spherical polars)

$$ds^2 = dr^2 + \sinh^2 r(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{58}$$

Thus r now measures geodesic distance from the origin; this coordinate system is known as geodesic polars.

The advantage of the spherical polars as such is that they simplify the action of the Killing vector field J_{XY}. In a similar way it is often useful to tailor-make a coordinate system so that one coordinate direction lies along some interesting vector field. As an example, consider the Killing vector field J_{XU}. Set

$$X = \sqrt{1+Y^2+Z^2}\sinh u \qquad U = \sqrt{1+Y^2+Z^2}\cosh u \tag{59}$$

and use (u, Y, Z) as intrinsic coordinates. Then

$$\partial_u = \frac{\partial X}{\partial u}\partial_X + \frac{\partial U}{\partial u}\partial_U = U\partial_X + X\partial_U = J_{XU} \ . \tag{60}$$

The coordinate u now lies along the Killing field.

Conformal compactification

Equipped with these ways of drawing pictures we can now begin to discuss the hyperbolic geometry itself. Let us begin by studying the conformal boundary. We discussed the conformal compactification of quadric surfaces in the previous chapter, but now we will approach it in a different manner that is applicable to much more general spaces. The basic idea is to introduce a new "unphysical" metric $\hat{g}_{ab}$ related to the true metric by a conformal rescaling:

$$d\hat{s}^2 = \hat{g}_{ab}dx^a dx^b = \Omega^2 g_{ab}dx^a dx^b = \Omega^2 ds^2 \ . \tag{61}$$

The conformal factor Ω is to be chosen such that infinity lies at a finite distance from any point in the interior of our space when measured with the unphysical metric. We can then add a boundary at infinity to our space and use the unphysicical metric to study asymptotic behaviour by means of ordinary differential geometry; the conformal compactification of a given space is a manifold-with-boundary whose interior is the original manifold. The physical metric is defined only on the interior while the unphysical metric is defined everywhere, which means that the conformal factor that relates them must vanish at the boundary while it must be non-vanishing throughout the interior.

The conformal compactification of $\mathbf{H}^3$ is easily obtained if we work with stereographic coordinates. We choose

$$\Omega = \frac{1-\rho^2}{2} \quad \Rightarrow \quad d\hat{s}^2 = dx^2 + dy^2 + dz^2 \ . \tag{62}$$

The unphysical metric becomes the ordinary flat metric, and the conformal boundary of $\mathbf{H}^3$ is the sphere at $\rho = 1$. The conformally compactified space is a closed ball in $\mathbf{E}^3$. If we use the half space representation instead the conformal boundary appears as an infinite plane; in this picture one point of the conformal boundary is missing.

Note that there is some arbitrariness involved here since we could multiply the conformal factor with any function that is well behaved at the boundary. This means that we are free to change the metric induced on the boundary by a conformal rescaling—the conformal structure on the boundary is a well defined concept, but its geometry is not. Apart from this ambiguity the nature of the conformal boundary is an intrinsic property of the geometry. As an illustration, let us show that the conformal boundary of ordinary flat space is necessarily a point. Using spherical polars, we are looking for a function Ω of r such that the distance from the origin (say) to infinity is finite:

$$\int_\gamma d\hat{s} = \int_\gamma \Omega ds = \int_0^\infty \Omega(r) dr < \infty \ . \tag{63}$$

This means that the function must fall to zero faster than one over r as r goes to infinity. But it then follows that the circumference of a circle around the origin as measured by the unphysical metric obeys

$$\lim_{r \to \infty} 2\pi\Omega(r)r = 0 \ . \tag{64}$$

Therefore the conformal boundary is a point. In effect we have shown that the conformal compactification of flat space is a sphere, while the conformal compactification of hyperbolic space is a ball. The conformal boundary is a point in the former case and a sphere in the latter. The moral of this is that in some sense there is a lot of space at infinity in a hyperbolic space.

Möbius transformations

If we restrict ourselves to two dimensions we can use a complex coordinate $z = x+iy$ and rely on known facts from complex analysis; in the end complex analysis and hyperbolic geometry will turn out to illuminate each other. Notably we will use Möbius transformations, whose properties I presume to be more or less known. A Möbius transformation is a transformation of the form

$$z \to z' = \frac{\alpha z + \beta}{\gamma z + \delta} \ , \qquad \alpha\delta - \beta\gamma \neq 0 \ . \tag{65}$$

It can be shown that this is the most general analytic one-to-one transformation of the compactified complex plane to itself, and it has the property that it maps circles to circles (or straight lines) and preserves angles. In other words it preserves the conformal structure on the Riemann sphere. One can show that it is possible to find a Möbius transformation that throws three arbitrary points on another triplet of arbitrary points, and that this requirement completely determines the transformation. The map of some fourth point follows from the fact that the harmonic cross ratio

$$\{z_1, b_2; z_2, b_1\} \equiv \frac{z_1 - b_2}{z_1 - b_1} \frac{z_2 - b_1}{z_2 - b_2} \tag{66}$$

is invariant under Möbius transformations.

The coordinate transformation that relates the Poincaré disk to the half space representation has to be a Möbius transformation since both of these coordinate systems show angles correctly. In stereographic coordinates the metric is

$$ds^2 = \frac{4}{(1-|z|^2)^2} dz d\bar{z} \ . \tag{67}$$

We can perform a Möbius transformation that takes

$$0 \rightarrow i \qquad 1 \rightarrow 0 \qquad \infty \rightarrow -i \ . \tag{68}$$

It is given by

$$z' = \frac{-iz + i}{z + 1} \ . \tag{69}$$

At this point it is useful to observe that

$$\alpha\delta - \beta\gamma = 1 \qquad \Rightarrow \qquad \frac{dz'}{dz} = \frac{1}{(\gamma z + \delta)^2} \ . \tag{70}$$

The normalization can always be arranged. It is now straightforward to express the metric in the new coordinates. If we revert to real coordinates $z' = x + iy$ at the end we find

$$ds^2 = \frac{1}{y^2}(dx^2 + dy^2) \ . \tag{71}$$

This is Poincaré's upper half plane picture. The Möbius transformation has transformed the conformal boundary of $\mathbf{H}^2$ from the unit circle to the real line.

The group of Möbius transformations is doubly covered by the group $SL(2, \mathbf{C})$, which consists of all two by two matrices of the form

$$G = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} , \qquad \alpha\delta - \beta\gamma = 1 \ . \tag{72}$$

These matrices act on the space of two component complex valued spinors, that can be regarded as homogeneous coordinates for $\mathbf{CP}^1 = \mathbf{S}^2$. We will need some quite detailed information about how Möbius transformations act, but it will be enough to understand three special families of such transformations having the property that the trace of their corresponding $SL(2, \mathbf{C})$ matrix is real. They are:

$$\text{Elliptic: } |\text{Tr}G| < 2 \qquad \text{Hyperbolic: } |\text{Tr}G| > 2 \qquad \text{Parabolic: } |\text{Tr}G| = 2 \ . \tag{73}$$

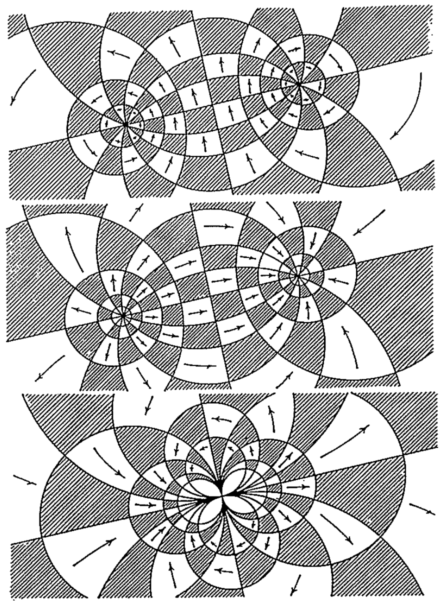

Figure 5: Elliptic, hyperbolic, and parabolic Möbius transformation. This unsurpassed illustration is taken from Lester Ford's book on *Automorphic functions.*

An elliptic Möbius transformation has two elliptic fixed points, a hyperbolic transformation has two hyperbolic fixed points and a parabolic one has a single fixed point that can be regarded as arising from the merger of two hyperbolic (say) fixed points. In all cases the flowlines are arcs of circles, and so are the level surfaces.

Non-Euclidean geometry

Let us now home in on the geometry of the hyperbolic plane. We know what the geodesics are in embedding coordinates, and it is therefore easy to show that on the Poincaré disk they are arcs of circles that are orthogonal to the boundary of the disk. The geodesic distance between two arbitrary points can be computed in the following manner: First choose one point at the origin and the other at radius r, and calculate

$$d = \int_0^\rho ds = \ln \frac{1+\rho}{1-\rho} \ . \tag{74}$$

Suppose that the geodesic cuts the boundary at $b_1 = -1$ and $b_2 = 1$. Then we can express this result as

$$d = \ln \frac{0-b_2}{0-b_1}\frac{z-b_1}{z-b_2} = \ln\{0, b_2; z, b_1\} \ . \tag{75}$$

The right hand is the logarithm of the harmonic cross ratio of the four points, which is known to be an invariant under Möbius transformations. Now it is clear that the isometries of $\mathbf{H}^2$ must be Möbius transformations (since these are the most general angle preserving transformations that provide one-to-one maps of the disk onto itself) and we know that $\mathbf{H}^2$ is a homogeneous space. Hence we can employ the isometry group to move the geodesic to arbitrary position without changing the value of the cross ratio, and express the geodesic distance as

$$d = \ln \frac{z_1-b_2}{z_1-b_1}\frac{z_2-b_1}{z_2-b_2} \ . \tag{76}$$

The Poincaré disk offers a concrete model of non-Euclidean geometry, the "new world" discovered at the beginning of the nineteenth century. The straight lines in this geometry are to be represented by the geodesics on the disk. A pair of points clearly define a unique geodesic, and if we run through the list of Euclid's axioms we see that they all are obeyed—with the exception of the fifth, since it is clear that through a point outside a given line we can now draw an infinite set of straight lines that are parallells in the sense that they do not intersect the given line. Finally, note that the geodesics tend to diverge from each other—taking the point of view of dynamical systems we might say that their Lyapunov exponent is positive, and indeed this turns out to be a fruitful point of view that pays a dividend on compact hyperbolic spaces.

To get a feeling for non-Euclidean geometry we consider some simple geometric figures. First we consider a circle with center at the origin. We use geodesic polar coordinates so that the metric is

$$ds^2 = dr^2 + \sinh^2 r d\phi^2 \ . \tag{77}$$

If the radius of the circle is R then the circumference C and the area A of the circle are

$$C = \int_0^{2\pi} \sinh R d\phi = 2\pi \sinh R \tag{78}$$

$$A = \int_A dA = \int_0^R \int_0^{2\pi} dr d\phi \sinh r = 2\pi(\cosh R - 1) \ . \tag{79}$$

Both of them grow exponentially with the radius. This reflects the fact that there is a lot of space in a negatively curved space.

The area of a polygon (bounded by segments of geodesics) is also instructive to compute. We choose to work in the upper half plane. Using Green's theorem

$$\int_A \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy = \int_{\partial A} u dx + v dy \tag{80}$$

we find that the area is

$$A = \int_A \frac{dx dy}{y^2} = \int_{\partial A} \frac{dx}{y} \tag{81}$$

where the integration is along the boundary of the polygon. Suppose that the polygon has n sides. To compute the contribution to the integral from a particular side we choose polar coordinates r, θ adapted to that particular segment; the range of θ is then from β to γ. Each segment will therefore give a contribution to the line integral which is

$$\int \frac{dx}{y} = \int \frac{d(r\cos\theta)}{r \sin\theta} = -\int_{\beta_i}^{\gamma_i} d\theta = \beta_i - \gamma_i \ . \tag{82}$$

Hence

$$A = \sum_{i=1}^{n} (\beta_i - \gamma_i) \ . \tag{83}$$

This formula is interesting because it can be used to relate the area to the sum of angles α_i at the vertices. Looking at the normal to the edges as it moves around the polygon, we see that its total rotation (2π) can be decomposed as

$$2\pi = \sum_i (\pi - \alpha_i) + \sum_i (\gamma_i - \beta_i) = n\pi - \sum_i \alpha_i - A \ . \tag{84}$$

The conclusion is that

$$A = (n-2)\pi - \sum_i \alpha_i \ . \tag{85}$$

As the polygon grows its angles have to become more acute. There is also an upper limit on how large the polygon can become. This is all very different from the scale invariant Euclidean geometry—the area does not enter the formula in the Euclidean case because the only change in the direction of the normal in a flat space takes place at the vertices, so that the angle sum in the Euclidean case is simply given by $(n-2)\pi$. The angle sum for a triangle in the hyperbolic plane on the other hand is always less than π; this is indeed not Euclidean geometry.

Isometries

Our next topic is isometries, or Killing vectors if we take the infinitesimal point of view. Since an isometry is a one-to-one map of the space to itself that preserves angles it must be a Möbius transformation, but it must be a restricted kind of Möbius transformation that maps the conformal boundary to itself; on the Poincaré disk the isometry group is the group of Möbius transformations that map the unit circle to itself and its interior to itself. They are of the form

$$z \to z' = \frac{\alpha z + \beta}{\bar{\beta} z + \bar{\alpha}} \ . \tag{86}$$

Strictly speaking one should add reflections ($z \to \bar{z}$) to this, but as a rule we will always ignore such discrete transformations. In the upper half plane the same group appears as the group of Möbius transformations that map the real line to itself and preserves the upper half plane. These are precisely the Möbius transformations with real coefficients. The double covering group in the first case is $SU(1,1)$, by definition the group of two by two matrices of the form

$$G \in SU(1,1) \quad \Rightarrow \quad G = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \qquad |\alpha|^2 - |\beta|^2 = 1 \ . \tag{87}$$

In the second case it is the group $SL(2, \mathbf{R})$ of real two-by-two matrices of unit determinant,

$$G \in SL(2, \mathbf{R}) \quad \Rightarrow \quad G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} , \qquad ad - bc = 1 \ . \tag{88}$$

These two groups are isomorphic; indeed since we already know from the hyperboloid picture that the isometry group of $\mathbf{H}^2$ is $SO_0(2,1)$ we have just sketched a proof that

$$SO_0(2,1) = SU(1,1)/\mathbf{Z}_2 = SL(2,\mathbf{R})/\mathbf{Z}_2 \ . \tag{89}$$

We wish to understand the action of these isometries in some detail.

It will be enough to understand one representative from each conjugacy class of the group. In general two elements g_0 and g_0' of a group are said to belong to the same conjugacy class if there exists a group element g such that

$$g_0' = g g_0 g^{-1} \ . \tag{90}$$

Thus in the case of $SO(3)$ Euler's theorem informs us that there is a one parameter family of rotations all of which have a given line of fixed points serving as the axis of rotation, and any rotation can be brought to this form by conjugation. Our group is a little more complicated: There are three families of conjugacy classes of isometries—since the trace of an $SU(1,1)$ or an $SL(2,\mathbf{R})$ matrix is automatically real they are elliptic, hyperbolic and parabolic Möbius transformations, respectively. After a brief investigation one finds that an elliptic isometry has a single elliptic fixed inside the disk—clearly they are analogous to rotations in $\mathbf{E}^2$. A hyperbolic isometry has two hyperbolic fixed points on the boundary of the disk—they are analogous to translations in $\mathbf{E}^2$, but unlike the latter they do not commute. They are sometimes called transvections. An important point to notice is that there is one and only one flow line of a transvection that is also a geodesic. Conversely, any geodesic determines a one parameter family of transvections. A parabolic isometry has one parabolic fixed point on the boundary of the disk, and its flow lines lie on spheres that just touch the boundary in a point. Such circles are called horocycles. Since the fixed point of an arbitrary rotation can be brought to the origin by means of a global isometry there is a one parameter family of conjugacy classes of elliptic isometries. Similarly there is a one parameter family of inequivalent hyperbolic isometries. At first sight one might be inclined to think that there should be a one-parameter family of inequivalent parabolic Möbius transformations as well, but this is not so since two parabolic Möbius transformations g_0 and g_o' sharing the same fixed point can be transformed into each other by choosing a suitable transvection g so that $g_0' = g g_0 g^{-1}$.

It is perhaps worth observing that the fact that the fixed points of the hyperbolic isometries lie on the boundary of the disk could have been predicted without any calculations: It is known that in the neighbourhood of a fixed point of an isometry of a space with a positive definite metric the isometry always looks like a rotation, so that only elliptic fixed points can occur in the interior. Continuing this argument we find—since the unphysical metric extends to the boundary—that a hyperbolic Möbius transformation must be accompanied by a conformal rescaling of the unphysical metric.

Hyperbolic three-space

The preceding discussion of the geometry of the Poincaré disk generalises to the Poincaré ball with little ado. Geodesics are arcs of circles that are orthogonal to the boundary. There will also be totally geodesic surfaces; intersections of the hyperboloid with timelike planes through the origin if you will. These are segments of spheres orthogonal to the boundary. Their intrinsic geometry is that of hyperbolic planes with the same radius of curvature as the three-space itself. The horospheres—spheres that touch the boundary in a single point—are flat. This statement is particularly easy to check if we use the half space representation, where the horospheres touching the point at infinity are represented by planes at constant coordinate distance from the boundary. A sphere lying entirely inside the Poincaré ball is what it looks like—a sphere of constant positive curvature. The properties of polyhedra are similar to those of polygons in the disk—if a polyhedron grows its angles must shrink.

We know from the hyperboloid picture that the isometry group is the

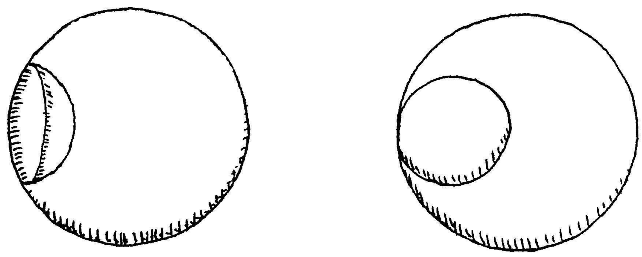

Figure 6: A totally geodesic surface and a horosphere.

Lorentz group $SO(3,1)$. On the other hand the conformal boundary is a two-sphere with a fixed conformal structure, and a significant part of the geometry can be controlled (as it were) from there. In particular the isometry group necessarily maps the boundary to itself and it must preserve the structure that is present there. This means that the isometry group is realised as the group of conformal transformations of the compactified complex plane, and this is precisely the group of unrestricted Möbius transformations. In outline this is the proof of the isomorphism

$$SO_0(3,1) = SL(2, \mathbf{C})/\mathbf{Z}_2 \ . \tag{91}$$

In this correspondence the rotations (such as the group elements generated by the Killing vectors J_{XY}, J_{YZ} and J_{ZX}) appear as elliptic Möbius transformations with two elliptic fixed points that in fact occur where an axis of rotations cuts the boundary. Group elements generated by boosts (like J_{XU}, J_{YU} and J_{ZU}) correspond to hyperbolic Möbius transformations with two hyperbolic fixed points on the boundary—in the interior they are transvections without fixed points, and their flow lines lie on segments of circles going through the fixed points. The level surfaces are segments of spheres, and there is precisely one flow line that is also a geodesic. The parabolic Möbius transformations with one fixed point on the boundary correspond to null rotations such as $J_{XY} + J_{UY}$; the flow in the interior takes place on horocycles and the level surfaces are again spheres. The half space representation is particularly well suited to study null rotations; choosing this representation so that the fixed point lies at infinity and its horospheres are planes at constant x we find that the metric is

$$ds^2 = \frac{1}{x^2}(dx^2 + dy^2 + dz^2) \tag{92}$$

and the Killing vector fields giving rise to null rotations are

$$J_{XY} + J_{UY} = \partial_y \qquad J_{XZ} + J_{UZ} = \partial_z . \tag{93}$$

So the null rotations take place in horospheres and in the half space representation they look just like translations of flat planes.

Horospheres and horocycles

Why are horospheres, or horocycles in the two dimensional case, important? The answer has to do with analysis. In flat space a plane is important—*inter alia*—because it can be a surface of constant phase for a plane wave, and plane waves are the basic objects of harmonic analysis. The same is true of horospheres in hyperbolic space.

Let us consider Helmholtz' equation

$$-\triangle\psi = \lambda\psi \ . \tag{94}$$

on the Poincaré disk, and start looking for plane wave solutions. What is a plane wave supposed to be? In flat space it can be characterized as a wave that propagates in the direction of a geodesic and has constant phase along planes that are normal to this geodesic. This suggests that we should look for solutions that are constant on horocycles, since horocycles can be regarded as planes that are orthogonal to a family of geodesic emerging from a source (or a sink) at infinity. The calculation is best made in the upper half plane, where the equation to be solved is

$$y^2(\partial_x^2 + \partial_y^2)\psi = -\lambda\psi \tag{95}$$

and the Ansatz is that ψ shall be independent of x. A set of solutions are evidently

$$\psi = y^s \qquad \Rightarrow \qquad \lambda = -s(s-1) \ . \tag{96}$$

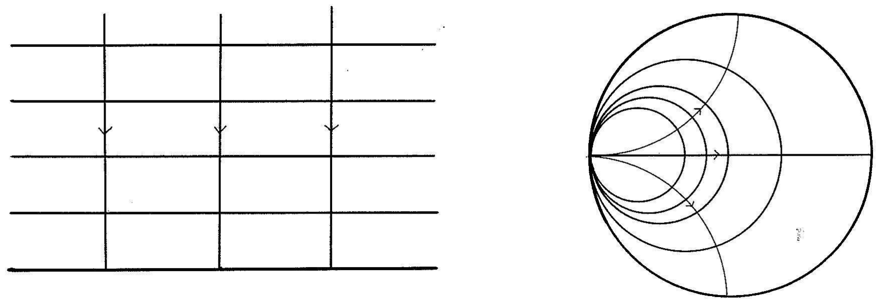

Figure 7: Plane waves: in the upper half plane and in the Poincaré disk.

The eigenvalues should be real, but if we choose s to be real we do not get travelling waves—in fact we get solutions that grow exponentially as they approach the boundary. The other possibility is

$$s = \frac{1}{2} + i\kappa \qquad \Rightarrow \qquad \lambda = \frac{1}{4} + \kappa^2 \ . \tag{97}$$

The spectrum is bounded from below. Since zero is not an eigenvalue there is also a "mass gap" in the spectrum. This is in fact what one might expect from hyperbolic geometry. At large length scales there is "more space" present than one would encounter in flat space. Therefore the infrared properties of a Green function may be expected to resemble the infrared properties of a Green function in some higher dimension—that is to say, it should be less singular in the infrared than a flat space Green function would be.

Anyway we can now rewrite our putative plane wave as

$$\psi = e^{(\frac{1}{2}+i\kappa)\ln y} \ . \tag{98}$$

But $\ln y$ is just the geodesic distance

$$d = \int ds = \int_1^y \frac{dy}{y} \tag{99}$$

from a point on a horocycle through $z = i$ to the horocycle at y. It follows that a plane wave can be written invariantly as

$$\psi = e^{(\frac{1}{2}+i\kappa)d} \ . \tag{100}$$

In flat space plane waves are important because of harmonic analysis—any reasonable function can be expressed as a linear combination of plane waves. *Mutatis mutandis* this is true in hyperbolic space as well, but for the moment we will drop the subject here.

Exercises:

- Perform the projection of the hyperboloid that gives the Klein ball, and verify that the metric takes the form stated.
- Show that a Möbius transformation may be uniquely "lifted" from the boundary to the interior of $\mathbf{H}^3$, and give explicit formulæ using (say) the half space picture.

ANTI-DE SITTER SPACE

The cosmological constant

We are now ready to discuss anti-de Sitter space in more detail. A first remark is that since it has constant negative curvature it solves Einstein's equations

$$R_{\alpha\beta} = \lambda g_{\alpha\beta} \tag{101}$$

with a negative cosmological constant. In the four dimensional case, when the quadric is

$$X^2 + Y^2 + Z^2 - U^2 - V^2 = -1 \ , \tag{102}$$

the cosmological constant has to take the value $\lambda = -3$ to ensure that Einstein's equations hold on the quadric. This is to say that we will use λ to define a unit of length, and therefore λ will not appear explicitly in our equations. That the cosmological constant provides spacetime with a natural length scale made a deep impression on Eddington. Echoing Gauss, he claimed that "to set the cosmical constant to zero would knock the bottom out of space". Physically speaking a negative cosmological constant corresponds in the Newtonian limit to an extra attractive term in the gravitational force. This is a useful fact to keep in mind.

Sausage coordinates

First we introduce what we call the sausage coordinates, in terms of which anti-de Sitter space will appear as a salami whose slices are hyperbolic spaces. Note how easy it is to slice a space of constant curvature with lower dimensional spaces of constant curvature, when the embedding coordinates are used in an intelligent fashion: Set

$$U = R\cot t \qquad\qquad V = R\sin t \ . \tag{103}$$

The quadric is then given by

$$X^2 + Y^2 + Z^2 - R^2 = -1 \ . \tag{104}$$

The anti-de Sitter metric becomes

$$ds^2 = dX^2 + dY^2 + dZ^2 - dR^2 - R^2 dt^2 \ . \tag{105}$$

For constant t, these are just the equations that define hyperbolic three-space as a quadric embedded in four dimensional Minkowski space. If—one way or the other—we introduce intrinsic coordinates on hyperbolic three-space, and denote its intrinsic metric by $d\sigma^2$, the anti-de Sitter metric is

$$ds^2 = -R^2 dt^2 + d\sigma^2 \ , \tag{106}$$

where R is some definite function of the intrinsic coordinates on hyperbolic three-space. Since R does not depend on t this is a static metric. The time coordinate t is periodic, which implies the presence of closed timelike curves.

We cannot draw a picture until we have introduced intrinsic coordinates on $\mathbf{H}^3$ as well. We will use stereographic coordinates (and spherical polars) for this purpose, so we set

$$X = \frac{2\rho}{1-\rho^2} \sin\theta \cos\phi \tag{107}$$

$$Y = \frac{2\rho}{1-\rho^2} \sin\theta \sin\phi \tag{108}$$

$$Z = \frac{2\rho}{1-\rho^2} \cos\theta \tag{109}$$

$$U = \frac{1+\rho^2}{1-\rho^2} \cos t \qquad V = \frac{1+\rho^2}{1-\rho^2} \sin t \ . \tag{110}$$

The angular coordinates have their usual range, while $0 \leq \rho < 1$. (If we set $\theta = \pi/2$ we obtain coordinates for three dimensional anti-de Sitter space sliced with hyperbolic planes.) We are now able to enjoy the two advantages that spatial infinity lies at a finite coordinate distance ($\rho = 1$) from the origin, which is fortunate when one wants to draw pictures, and that the metric on the spatial slices is manifestly conformally flat, so that all spatial angles will be faithfully represented by the pictures. Indeed the anti-de Sitter metric in sausage coordinates becomes

$$ds^2 = -\left(\frac{1+\rho^2}{1-\rho^2}\right)^2 dt^2 + \frac{4}{(1-\rho^2)^2}(d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2\theta d\phi^2) \ . \qquad (111)$$

So now we can draw an "intrinsic" picture of three-dimensional anti-de Sitter space (set $\theta = \pi/2$ in the formulas). Note that the surfaces $t = 0$ and $t = 2\pi$ have to be identified if the picture really is to depict anti-de Sitter space. "Going to the covering space" means that this is ignored, and the cylinder is continued indefinitely in both directions. There are many contexts in which the distinction between anti-de Sitter space and its universal covering space is immaterial, and sometimes I permit myself to choose one or the other without explicitly saying so.

Anti-de Sitter space as a sausage sliced with hyperbolic planes

Other coordinates

Personally I prefer to stick to sausage coordinates for visualisation and embedding coordinates for calculations, but the whole range of coordinate systems for hyperbolic space—including half space coordinates, stereographic coordinates, and all sorts of *ad hoc* concoctions—have their analogues in anti-de Sitter space. The stereographic coordinates have the advantage that lightcones look like lightcones. This is so because the metric becomes manifestly conformally flat, and the path of a null geodesic does not depend on the conformal factor of the metric.

The details are as follows: The projection will be made from the point $(X, Y, Z, U, V) = (0, 0, 0, 0, -1)$ onto the plane at $V = 0$. Since a particular choice of steregraphic coordinates is singled out by choosing the point antipodal to the point of projection we describe this choice as being "centered on $V = 1$"; around this point distances are not as distorted by the projection as elsewhere. The region covered by the projection is $V > -1$, or in words it is the interior of the light cone at the point of projection. (The fact that this coordinate system does not cover spacetime globally detracts from its usefulness.) Explicitly

Stereographic projection of anti-de Sitter space

$$X = \frac{2x}{1-s^2} \qquad Y = \frac{2y}{1-s^2} \qquad Z = \frac{2z}{1-s^2} \tag{112}$$

$$U = \frac{2u}{1-s^2} \qquad V = \frac{1+s^2}{1-s^2} \ , \tag{113}$$

where

$$s^2 \equiv x^2 + y^2 + z^2 - u^2 < 1 \ . \tag{114}$$

The intrinsic metric in these coordinates is manifestly conformally flat, as advertised:

$$ds^2 = \frac{4}{(1-s^2)^2}(dx^2 + dy^2 + dz^2 - du^2) \ . \tag{115}$$

The coordinate space is now that subset of Minkowski space which is contained within the one-sheeted hyperboloid defined by $s^2 = 1$. The waist of this hyperboloid is a Poincaré ball described by the embedding coordinates as $U = 0$; the point $V = 1$ sits at the center of the waist. The surface $V = 0$ becomes a hyperboloid in the flat coordinate space, making it manifest that its intrinsic geometry is that of $\mathbf{H}^3$.

Anti-de Sitter space in stereographic coordinates

Conformal compactification and $\mathscr{J}$

The pictures that we drew really depict conformally compactified anti-de Sitter space; we select a conformal factor Ω so that the conformally related metric is well behaved on the boundary $\rho = 1$. As always the precise choice of this factor is a matter of judgment. A natural choice is

$$\Omega = \frac{1-\rho^2}{1+\rho^2} = \frac{1}{\sqrt{U^2+V^2}} \ . \tag{116}$$

Then (using sausage coordinates)

$$d\hat{s}^2 = \Omega^2 ds^2 = -dt^2 + \frac{4}{(1+\rho^2)^2}(d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2\theta d\phi^2) \ . \tag{117}$$

The spatial metric is that of the three-sphere rather than flat space. Indeed the unphysical spacetime metric is that of the Einstein universe, which has the topology $\mathbf{S}^3 \otimes \mathbf{R}$, and what we have shown is that anti-de Sitter space is conformally related to a subset of the Einstein universe, as depicted in the figure (where two dimensions have been suppressed).

Anti-de Sitter space in the Einstein universe

The boundary of conformally compactified $\mathbf{adS}_4$ has the topology $\mathbf{S}^2 \otimes \mathbf{R}$, where the sphere can be regarded as the conformal boundary of hyperbolic three-space. This boundary is timelike. It is common practice to refer to it as $\mathscr{J}$ or scri—"script I"—which is defined as the set of endpoints of all future directed (or past directed, as the case may be) lightlike geodesics. Of course the boundary is also the set of endpoints of spatial geodesics, so we can alternatively refer to it as spatial infinity, but lightlike geodesics are more important for the causal structure. The whole structure is quite different from that of conformally compactified Minkowski space, for which spatial infinity, future $\mathscr{J}$, and past $\mathscr{J}$ are disjoint, and the latter two are lightlike. It is a crucial difference and we will devote a chapter of its own to its study.

Cauchy developments

We now try to put some more structure into our picture of the interior. To do this we solve the equations for a lightlike geodesic moving radially outwards from the origin:

$$ds^2 = 0;\ d\theta = d\phi = 0 \quad \Rightarrow \quad dt = \pm\frac{2d\rho}{1+\rho^2} \quad \Rightarrow \quad \tan\frac{t}{2} = \pm\rho\ . \tag{118}$$

The conclusion is that a light ray that leaves the origin at $t = 0$ ends on $\mathscr{J}$ at $t = \pi/2$. This is one of the results that we have already derived using the embedding coordinates; another of those results imply that radially directed timelike geodesics that start out at the origin eventually turn around and reconverge at $r = 0$ after the passage of an amount π of time t. This makes physical sense, and reflects the fact that a negative cosmological constant corresponds to an attractive gravitational force. An apparent drawback of our sausage coordinates is that the slope of a light ray depends on the radius. We can rescale the radial coordinate so that all radial lightlike geodesics get slope one (as one does when drawing Penrose diagrams), but then it turns out that the slope depends on direction instead. Our choice of radial coordinate ensures that the slope of a light ray is independent of direction, although it does depend on the radius.

Lightlike, timelike and spacelike geodesics in anti-de Sitter space

An important property of anti-de Sitter space is that it is not globally hyperbolic. What this means is that there are no Cauchy hypersurfaces in this spacetime, since information is always "leaking in" from its timelike boundary. Given initial data on (say) the hyperbolic three-space defined by $t = 0$, we cannot predict all of the future, and indeed after the passage of an amount $\pi/2$ of time we have completely lost control of the time development, unless we can somehow control the influx of information from infinity. We can introduce a new coordinate system to emphasize this state of affairs. It will cover only the Cauchy development of the surface $V = t = 0$; the idea is to rewrite the quadric that defines **adS** as

$$X^2 + Y^2 + Z^2 - U^2 = -(1 - V^2) \ . \tag{119}$$

Provided that $|V| < 1$ the surfaces of constant V are hyperbolic three-spaces with V-dependent curvature. But the hypersurface $V = 1$ is the backwards lightcone of the point at the spatial origin at $t = \pi/2$—this is manifest in our stereographic coordinates where it becomes a lightcone in coordinate space with its vertex at the center of the waist—and it hits the boundary at $V = 0$. Hence $|V| < 1$ is precisely the Cauchy development of the surface $V = 0$; the conclusion is that the Cauchy development can be foliated with hyperbolic three-spaces of growing intrinsic curvatures.

The Cauchy development of $t = 0$ sliced with hyperbolic spaces

Let us be a little more explicit about this; we set

$$V = \sin T \ . \tag{120}$$

We rescale the remaining coordinates according to

$$X = \hat{X}\cos T \qquad Y = \hat{Y}\cos T \qquad Z = \hat{Z}\cos T \qquad U = \hat{U}\cos T \ , \quad (121)$$

Then the defining equation for anti-de Sitter space is obeyed provided that

$$\hat{X}^2 + \hat{Y}^2 + \hat{Z}^2 - \hat{U}^2 = -1 \ . \quad (122)$$

This equation defines a hyperbolic three-space. The anti-de Sitter metric in these coordinates is

$$ds^2 = \cos^2 T(d\hat{X}^2 + d\hat{Y}^2 + d\hat{Z}^2 - d\hat{U}^2) - dT^2 = -dT^2 + \cos^2 T d\sigma^2 \ . \quad (123)$$

where $d\sigma^2$ is a metric on $\mathbf{H}^3$. This is a metric of the Robertson-Walker form; the Cauchy development has been foliated with hyperbolic three-spaces with constant negative curvatures that grow with T—indeed the intrinsic curvature of the spatial slices is proportional to $-1/\cos^2 T$. The curvature diverges when $T = \pi/2$. From the picture it is clear that this happens because the hyperboloids degenerate to a cone at this moment in time. The region outside the light cone where the "Cauchy coordinates" break down can be sliced with three-dimensional de Sitter spaces if need be.

Null planes

Once we know what a light ray looks like it is natural to ask for the anti-de Sitter analogue of a null plane. By definition, a null or lightlike surface is a surface that is contains its own normal. Since the normal is orthogonal to the surface it must be a lightlike vector, and it must be the only lightlike vector in the surface since a surface containing two null vectors is timelike. A null surface is in fact ruled by a set of null geodesics.

A lightcone is a null surface. A null plane can be thought of as a lightcone whose vertex sits at infinity. It is a totally geodesic surface and we know what those are in anti-de Sitter space, namely surfaces of the form

$$a \cdot X = 0 \ . \quad (124)$$

This is a null surface if

$$a^2 = 0 \ . \tag{125}$$

It is useful to know how such a surface looks in sausage coordinates (although the corresponding formula is not so useful).

A null plane in sausage coordinates

It is also instructive to work out what it looks like in stereographic coordinates, where a lightcone looks like a lightcone. Taking out a factor if need be the equation can be written as

$$x_0X + y_0Y + z_0Z - u_0U - V = 0 \tag{126}$$

where

$$x_0^2 + y_0^2 + z_0^2 - u_0^2 = 1 \ . \tag{127}$$

But this means that (x_0, y_0, z_0, u_0) are the stereographic coordinates of some point on $\mathcal{J}$, and it is a minor exercise to show that the equation for the null plane can be rewritten in terms of stereographic coordinates as

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = (u - u_0)^2 \ . \tag{128}$$

This is indeed the equation for a lightcone with its vertex on $\mathcal{J}$, and confirms that we are indeed dealing with a null plane.

Optical geometry

Let us pursue the propagation of light in anti-de Sitter space a bit further. Since it is a static spacetime, we can identify all spaces of constant t and ask for the spatial trajectory of a light ray. In other words, we choose the manifestly static sausage coordinates and then we look straight down the tube from above—we will see the light rays projected onto a spatial disk. With the exception of those rays that go through the origin, the spatial paths followed by the light rays are not geodesics on the Poincaré disk (whereas in flat spacetime they would be straight lines in space). However, for all static spacetimes it is true that the light rays do follow spatial paths that are geodesics with respect to a spatial metric which is defined in a different way than the actual "physical" metric. The idea is as follows: Consider the static metric

$$ds^2 = g_{tt}dt^2 + d\sigma^2 \ , \tag{129}$$

where all the components are independent of t and $d\sigma^2$ is the physical metric on space (i.e. that metric which is induced on a spatial slice by the metric on spacetime). Then the optical metric on space is defined by

$$dl^2 = \frac{1}{g_{tt}}d\sigma^2 \ . \tag{130}$$

It is now possible to prove that the spatial trajectory $x^a(t)$ of the light ray—which is a null geodesic with respect to the spacetime metric—is indeed a geodesic with respect to the optical metric, with the time coordinate t as its affine parameter. Let us accept this theorem (which is easy to prove) as a fact. Using the sausage coordinates we find that the optical metric for $\mathbf{adS}_4$ is

$$dl^2 = \frac{4}{(1+\rho^2)^2}(d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2\theta d\phi^2) \ . \tag{131}$$

So while the physical metric is that of $\mathbf{H}^3$, the optical metric is the metric of $\mathbf{S}^3$, which has constant positive curvature. In both cases the range of the coordinate ρ is restricted to $0 < \rho < 1$, so that we have to imagine only one hemisphere of $\mathbf{S}^3$ in order to visualize the light rays as geodesics. For an observer sitting inside anti-de Sitter space, infinity "looks like" the (two dimensional) equator of the three sphere, as seen from its North Pole.

Equipped with these ideas we can return to consider a family of light rays coming from a point on the boundary of anti-de Sitter space, choosing $\mathbf{adS}_3$ as an illustrative example. If we think of space as a hemisphere, the rays follow great circles on the sphere. If we use stereographic coordinates x, y on the sphere these great circles correspond to arcs of circles in the coordinate disk, but they are not orthogonal to its boundary (except in the special case that the circle is a straight line). Therefore the trajectories are not geodesics with respect to the physical metric. The particular trajectory going through the origin is a geodesic with respect to the physical metric though; there is no conflict with homogeneity of space here since the optical metric is not left invariant by $SO(2,1)$, and therefore the trajectories do not transform in any obvious way under Möbius transformations.

Light rays projected on the Poincaré disk

Isometries

We have yet to discuss the isometries of anti-de Sitter space. In a later chapter we will give a very detailed discussion for the special case of $\mathbf{adS}_3$, but we give the most important features right away. The isometry group of $\mathbf{adS}_4$ is $SO(3,2)$ and it has ten generators. There is one Killing vector that acts like time translation in Minkowski space. In sausage coordinates we have

$$\partial_t = \frac{\partial U}{\partial t}\partial_U + \frac{\partial V}{\partial t}\partial_V = U\partial_V - V\partial_U = J_{UV} \ . \tag{132}$$

There is also an $SO(3)$ subgroup generated by the three Killing vectors J_{XY}, J_{YZ} and J_{ZX}. They work just like rotations in Minkowski space.

So far we are missing the analogues of spatial translations and Lorentz boosts. We have six generators left to work with, and we pick J_{XU} as a representative example. We know that it acts like a transvection (with two fixed points on $\mathscr{J}$) on the hyperbolic three-space defined by $V = 0$, but it does not stay that way. Its norm squared is

$$||J_{XU}||^2 = ||X\partial_U + U\partial_X||^2 = U^2 - X^2 = Y^2 + Z^2 - V^2 + 1 \ . \tag{133}$$

Hence it is spacelike within the Cauchy development $|V| < 1$ of the surface $V = 0$, but there is also a region where it is timelike. On the surface defined by $X = U = 0$ it has a plane of fixed points, and close to this plane it behaves just like a Lorentz boost in Minkowski space. Somehow what started out like a spatial translation at $t = 0$ has become like a Lorentz boost at $t = \pi/2$.

The flow of J_{XU}

For reference we give all the Killing vectors of $\mathbf{adS}_3$ in sausage coordinates:

$$J_{UV} = \partial_t \qquad\qquad J_{XY} = \partial_\phi \tag{134}$$

$$J_{XU} = -\frac{2\rho}{1+\rho^2}\sin t\cos\phi\partial_t + \frac{1-\rho^2}{2}\cos t\cos\phi\partial_\rho - \frac{1+\rho^2}{2\rho}\cos t\sin\phi\partial_\phi \tag{135}$$

$$J_{YU} = -\frac{2\rho}{1+\rho^2}\sin t\sin\phi\partial_t + \frac{1-\rho^2}{2}\cos t\sin\phi\partial_\rho + \frac{1+\rho^2}{2\rho}\cos t\cos\phi\partial_\phi \tag{136}$$

$$J_{XV} = \frac{2\rho}{1+\rho^2}\cos t\cos\phi\partial_t + \frac{1-\rho^2}{2}\sin t\cos\phi\partial_\rho - \frac{1+\rho^2}{2\rho}\sin t\sin\phi\partial_\phi \tag{137}$$

$$J_{YV} = \frac{2\rho}{1+\rho^2}\cos t\sin\phi\partial_t + \frac{1-\rho^2}{2}\sin t\sin\phi\partial_\rho + \frac{1+\rho^2}{2\rho}\sin t\cos\phi\partial_\phi \ . \tag{138}$$

We will return to a systematic study of the isometry group of $\mathbf{adS}_3$ in a later chapter.

Killing horizons

In Lorentzian geometry isometries have some features that are not present when the metric is positive definite. This can be seen by thinking about a Lorentz boost in Minkowski space. First of all it has hyperbolic fixed points, something that cannot occur in the positive definite case. There are two regions where the flow is timelike and two where it is spacelike, and these regions are separated by a bifurcate null surface—having two sheets that cross each other on the set of fixed points—whose normal is the Killing vector itself. This situation is generic for spacetime isometries and motivates a definition: A null surface whose normal is a Killing vector is called a Killing horizon.

The bifurcate Killing horizon of a Lorentz boost

Let the Killing vector be ξ^α. We can define a family of hypersurfaces by

$$\xi^\alpha \xi_\alpha = \text{constant} . \tag{139}$$

The Killing horizon occurs where the constant vanishes. Since the normal lies along the Killing vector field there must exist a proportionality constant κ such that

$$\nabla_\alpha \xi^\beta \xi_\beta = -2\kappa \xi_\alpha . \tag{140}$$

This equation must hold on the null surface if it is indeed a Killing horizon. The proportionality constant κ is called the surface gravity of the horizon—

and its value must be evaluated on the horizon itself. *A priori* the surface gravity might be a scalar field rather than a constant, but it can be shown that it is necessarily constant in the direction of the null generators. Moreover it is constant in all directions on any bifurcate Killing horizon. There are Killing horizons that are not bifurcate but consist of a single sheet. For them the constancy of κ can be shown to hold if Einstein's equations are assumed to hold, with some reasonable conditions on the energy-momentum tensor. The bifurcate case is the generic one and occurs whenever there is some $(D-2)$ dimensional set of fixed points in a D dimensional spacetime.

If the surface gravity vanishes the Killing horizon is said to be degenerate. The flow of the Killing vector has to change from spacelike to timelike across a non-degenerate Killing horizon, but this may or may not be true if the horizon is degenerate. Conversely it should be noted that the Killing vector can change from spacelike to timelike without going through a Killing horizon, since the surface where the Killing vector becomes lightlike can be timelike.

Anti-de Sitter space provides instructive examples of the behaviour of Killing horizons. Let us first consider the Killing vector

$$\xi = \xi^\alpha \partial_\alpha = aJ_{XU} \ , \tag{141}$$

where a is some arbitrary constant. This Killing vector becomes lightlike when

$$\xi^\alpha \xi_\alpha = ||aJ_{XU}||^2 = a^2(U^2 - X^2) = a^2(U+X)(U-X) = 0 \ . \tag{142}$$

This is indeed a bifurcate null surface consisting of two intersecting null planes. On that branch of the surface where $X = U$ we find that

$$\nabla_\alpha a^2(U^2 - X^2) = 2a^2(-X, 0, 0, U, 0) = 2a^2(-U, 0, 0, X, 0) = -2a\xi_\alpha \ . \tag{143}$$

Therefore this is a bifurcate Killing horizon with surface gravity $\kappa = a$. (On the other branch of the surface the sign is reversed since the Killing vector is past directed there.)

Now suppose that we twist our Killing vector a little. First set

$$\xi = aJ_{XU} + bJ_{YV} \ , \qquad a > b > 0. \tag{144}$$

Then the Killing vector becomes lightlike on the surface

$$\xi^\alpha \xi_\alpha = a^2(U^2 - X^2) + b^2(V^2 - Y^2) = (a^2 - b^2)(U^2 - X^2) + b^2 = 0 \ . \quad (145)$$

However, it is easy to convince oneself that this is a timelike surface. Therefore it is not a Killing horizon and indeed there are none for this choice of Killing vector. Parenthetically it is clear from this observation of an invariant property that the twisted Killing vector must belong to a different conjugacy class than does aJ_{XU}. Next we take

$$\xi = a(J_{XU} + J_{XY}) \ . \quad (146)$$

This is an interesting case because the norm squared vanishes when

$$||a(J_{XU} + J_{XY})||^2 = a^2(U + Y)^2 = 0 \ . \quad (147)$$

This is a null surface with a single branch. Since it has a double zero it is clear that on this surface

$$\nabla_\alpha (U + Y)^2 = 0 \ . \quad (148)$$

The surface gravity is zero; this is a degenerate Killing horizon and the flow of the Killing vector is spacelike on both sides of the horizon.

The surfaces where a Killing vector becomes null; three cases

The attractive gravitational force

We can get some physical feeling for anti-de Sitter space if we compare an observer moving along the Killing vector field J_{XV} with an observer moving along a boost Killing vector field in Minkowski space. (We choose J_{XV}

because we tend to regard the surface $V = 0 = t$ as if it were our country of origin in anti-de Sitter space—and it is J_{XV} rather than J_{XU} which is timelike there.) For simplicity we restrict ourselves to the two dimensional case,

$$X^2 - U^2 - V^2 = -1 \ . \tag{149}$$

We also restrict our attention to the region where $V^2 < X^2$, which is where J_{XV} is timelike (and where $U^2 > 1$). In general, the acceleration experienced by an observer moving along a vector field ξ^α is given by

$$a^\alpha = \frac{1}{(-\xi^\gamma \xi_\gamma)} \xi^\beta \nabla_\beta \xi^\alpha = \frac{1}{2} \nabla^\alpha \ln\left(-\xi^\gamma \xi_\gamma\right) \ , \tag{150}$$

where the second step was possible only because we assume that the trajectory lies along a Killing vector field. The result of a small calculation (it is convenient to do it in a coordinate system where $J_{XV} = \partial_\tau$ for some suitably chosen time coordinate τ) then reveals that

$$a^2 \equiv a^\alpha a_\alpha = \frac{U^2}{U^2 - 1} \qquad \Rightarrow \qquad 1 < a < \infty \ . \tag{151}$$

The expression diverges when we approach the Killing horizon. The existence of a lower bound greater than zero is a new feature not present in Minkowski space. We can understand why it occurs if we consider the acceleration of an observer moving along the Killing vector J_{UV}, which in sausage coordinates is just ∂_t; this is an observer hovering at a constant distance from the origin of the sausage coordinates. For such an observer we obtain

$$a^2 = \frac{X^2}{1 + X^2} \qquad \Rightarrow \qquad 0 < a < 1 \ . \tag{152}$$

This observer is subject to an attractive gravitational force not present in Minkowski space, and the same force is acting on the accelerated observer.

Euclidean section

In physical applications it is frequently interesting to analytically continue a metric of Lorentzian signature to one of Euclidean signature—assuming this can be done at all. What we require for this purpose is a complex manifold that has the given Lorentzian space as a real section, and additionally has a real section with Euclidean signature. In practice such an Euclidean section may or may not exist, and it may or may not be unique if it exists. This raises some difficulties, but in the case of static spacetimes there are none: An Euclidean section of a static spacetime can always be obtained by letting

$$t \rightarrow -it \tag{153}$$

for the coordinate along the timelike Killing field.

Anti-de Sitter space is a static spacetime, so let us perform this analytic continuation using sausage coordinates. The resulting metric is

$$ds^2 = \left(\frac{1+\rho^2}{1-\rho^2}\right) dt^2 + \frac{4}{(1-\rho^2)^2}(dx^2 + dy^2 + dz^2) \ . \tag{154}$$

It is not difficult to see that these coordinates parametrize hyperbolic space

$$X^2 + Y^2 + Z^2 + V^2 - U^2 = -1 \ . \tag{155}$$

The continuation entails

$$V \rightarrow -iV \ : \qquad U = \frac{1+\rho^2}{1-\rho^2}\cosh t \qquad V = \frac{1+\rho^2}{1-\rho^2}\sinh t \ . \tag{156}$$

Therefore we regard hyperbolic four-space as the Euclidean counterpart of anti-de Sitter space. In a similar way the four-sphere is the analytic continuation of de Sitter space.

The link between $\mathbf{adS}_4$ and $\mathbf{H}^4$ can be made a good deal tighter. To see how, let us first forge the corresponding link between $\mathbf{dS}_4$ and $\mathbf{S}^4$. The spacetime topology is that of $\mathbf{S}^3 \otimes \mathbf{R}$. Now choose a three-sphere in de Sitter space that has vanishing extrinsic curvature (say, the "waist" of the hyperboloid). The data that we are given is the extrinsic curvature and the intrinsic metric on the three-space, and as is well known these data suffice to reconstruct de Sitter space both to the future and to the past of the three-sphere by means of Einstein's equations. On the other hand we can regard the three-sphere as the equator of a four-sphere. Since the extrinsic curvature vanishes anyway this is a possible interpretation, and the intrinsic metric on the three-sphere can be regarded as boundary data from which we can reconstruct the four-sphere by means of the Euclidean version of Einstein's equations. Now the idea is to use the Lorentzian equations on one side of the three-sphere, and the Euclidean equations on the other. The result is a smooth space that changes signature across the three-sphere. Not only is there a complex manifold that admits both $\mathbf{dS}_4$ and $\mathbf{S}^4$ as real sections—now we see that these sections can be made to intersect along an $\mathbf{S}^3$ inside the complex manifold.

Gluing a Lorentzian and an Euclidean space together

The same argument applies to $\mathbf{adS}_4$ and $\mathbf{H}^4$; evidently they intersect in an $\mathbf{H}^3$ with vanishing extrinsic curvature. There is a slight complication because the space is open, and the data on $\mathbf{H}^3$ have to be supplemented with data along the conformal boundaries of the spacetime and the space that we are trying to reconstruct. But this does not change the overall picture. Why is the picture wanted in the first place? The answer is that this kind of construction appears in the Euclidean approach to quantum gravity and in quantum cosmology; maybe it has something to do with physics.

Exercises:

- Draw all the pictures for this chapter.

ASYMPTOTIA

A change of hats

To set the conformal compactification of anti-de Sitter space in context we will wish to compare it to that of Minkowski space on the one hand, and to that of the more general class of asymptotically anti-de Sitter spacetimes on the other. Because we will focus on the unphysical geometry in the neighbourhood of $\mathscr{J}$ it will be convenient to make an exchange of hats—in this chapter all geometrical objects built from the physical metric will appear with hats on, while there will be no hats on unphysical objects.

Minkowski space

Minkowski space appears to be rather simpler than anti-de Sitter space, but this is not so when we try to understand what goes on at infinity—the structure of the conformal boundary of Minkowski space is quite a bit more involved than that of anti-de Sitter space. To get to grips with the former we begin by introducing retarded and advanced null coordinates

$$u = t - r \qquad\qquad v = t + r \ . \tag{157}$$

An outgoing radially directed null geodesic has constant u, which is why u is called a retarded coordinate—the coordinate v serves as an affine parameter along the outgoing ray. Anyway, in these coordinates the physical Minkowski metric is

$$d\hat{s}^2 = -dudv + \frac{(u-v)^2}{4}d\Omega^2 \ , \tag{158}$$

where $d\Omega^2$ is the metric on the two-sphere. Next we bring infinity in to a finite coordinate distance through

$$u = \tan p \qquad\qquad v = \tan q \ . \tag{159}$$

Then the metric becomes

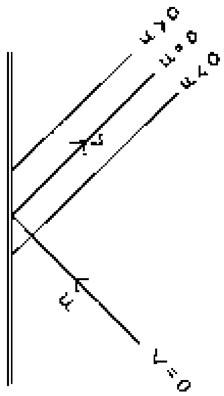

Figure 8: Advanced and retarded null coordinates.

$$d\hat{s}^2 = \frac{1}{4\cos^2 p \cos^2 q}(-4dpdq + \sin^2{(p-q)}d\Omega^2) \ . \tag{160}$$

"Infinity" now occurs at p or q equal to $\pi/2$, and the metric is clearly ill-defined there. This can be cured by a conformal rescaling; evidently we can choose the conformal factor so that the conformally related unphysical metric is

$$ds^2 = \Omega^2 d\hat{s}^2 = -4dpdq + \sin^2{(p-q)}d\Omega^2 \ . \tag{161}$$

(The conformal factor vanishes at the boundary and has non-zero gradient there, as it should.) This metric is perfectly well defined also when $p = \pi/2$. There is a coordinate singularity at $p = q$, but this does not matter. There is a similar problem at $(p, q) = (-\pi/2, \pi/2)$, corresponding to $r = \infty$. This is in fact another coordinate singularity and does not matter either. The point where it occurs is known as spatial infinity.

If we introduce new coordinates

$$\tau = p + q \qquad\qquad \rho = q - p \tag{162}$$

we get

$$ds^2 = -d\tau^2 + d\rho^2 + \sin^2\rho d\Omega^2 \ . \tag{163}$$

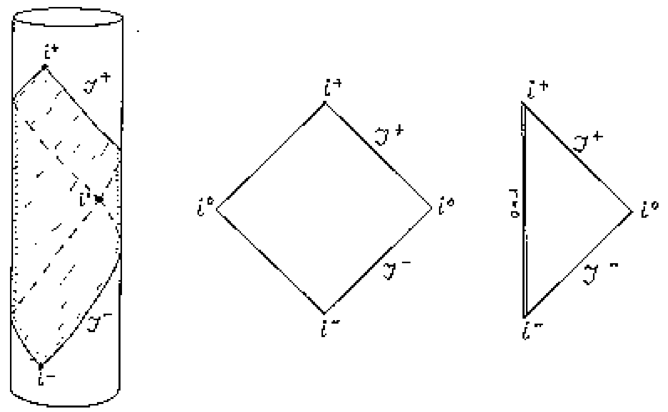

Figure 9: Here we see 1+1 dimensional Minkowski space, embedded in the Einstein universe, as a conformal diagram, and as a Carter-Penrose diagram. The latter is valid in any dimension.

This is the metric of the Einstein universe. The two coordinate singularities that we encountered above occur at antipodal points on the spatial three-sphere. So just like anti-de Sitter space Minkowski space is conformal to a subset of the Einstein universe, but the two subsets are different and the nature of the conformal boundary is dramatically different in the two cases. The boundary now consists of two separate null surfaces called future and past scri, $\mathscr{J}^+$ and $\mathscr{J}^-$. These are really lightcones with one vertex at future and past timelike infinity, i^+ and i^- respectively, and one vertex at spatial infinity i^0. The fact that a single light cone manages to have two vertices is a little hard to visualize directly, but it is clear from the way that it is embedded in the Einstein universe that this is what happens.

We can think of lightlike infinity $\mathscr{J}$ as the set of endpoints of lightlike geodesics, and spatial infinity i^0 as the set of endpoints of spacelike geodesics. It is the conformal rescaling that provides the geodesics with endpoints in the first place. In anti-de Sitter space the conformal boundary plays both roles, but in Minkowski space they are kept separate.

The picture that we have arrived at is not quite the same as the one that we would get from the point of view of the projective null quadric

$$X_1^2 + X_2^2 + X_3^2 + X_4^2 = X_5^2 + X_6^2 \tag{164}$$

in $\mathbf{RP}^5$. As we have seen this quadric provides a model for compactified Minkowski space from a conformal point of view. By scaling the homogeneous coordinates suitably we see that the topology of our quadric is $\mathbf{S}^3 \otimes \mathbf{S}^1$. We can recover this structure from the new version of compactified Minkowski space if we identify $\mathcal{J}^+$ and $\mathcal{J}^-$ with each other. This can be done in a fairly natural way since every null plane in Minkowski space can be regarded as a lightcone with one vertex on $\mathcal{J}^+$ and the other on $\mathcal{J}^-$; one can identify the corresponding points and in the process one finds that i^+, i^- and i^0 get identified with each other. In this way all lightrays become topological circles and there is a single lightcone at infinity. As a reward we gain a natural action of the conformal group—the drawback is that no such manœuvre is possible in more general spacetimes. Since Minkowski space is not our subject we do not enter into any further details here.

Conformally related spacetimes

As a preparation for our study of the conformal boundary of more general spacetimes we consider two conformally related but otherwise arbitrary spacetimes in some detail. Suppose that

$$g_{\alpha\beta} = \Omega^2 \hat{g}_{\alpha\beta} \ . \tag{165}$$

We introduce the rule that indices on hatted and unhatted objects are raised and lowered with the hatted and the unhatted metric, respectively. It follows that

$$\begin{gathered} T_{\alpha_1 \dots \alpha_m}{}^{\beta_1 \dots \beta_n} = \Omega^{w-n+m} \hat{T}_{\alpha_1 \dots \alpha_m}{}^{\beta_1 \dots \beta_n} \\ \Rightarrow \\ T_{\alpha_1 \dots \alpha_{m-1}}{}^{\alpha_m \beta_1 \dots \beta_n} = \Omega^{w-(n-1)+(m-1)} \hat{T}_{\alpha_1 \dots \alpha_{m-1}}{}^{\alpha_m \beta_1 \dots \beta_n} \ . \end{gathered} \tag{166}$$

The number w is known as the conformal weight of the tensor. The metric has weight zero, and the definition is made in such a way that the weight is unchanged by raising and lowering of indices. It is only rather special tensors that transform in this way—most tensors do not have a conformal weight.

We will want to remove the hats from various geometric objects. First we define

$$n_\alpha = \nabla_\alpha \Omega \ . \tag{167}$$

Then it is easy to show that

$$\hat{\nabla}_\alpha V_\beta = \nabla_\alpha V_\beta + C_{\alpha\beta}{}^\gamma V_\gamma \ , \tag{168}$$

where the contortion tensor is

$$C_{\alpha\beta}{}^\gamma = \frac{1}{\Omega}(\delta^\gamma_\alpha n_\beta + \delta^\gamma_\beta n_\alpha - g_{\alpha\beta} n^\gamma) \ . \tag{169}$$

It is a straightforward exercise to derive from this that

$$\hat{R}_{\alpha\beta}{}^{\gamma\delta} = \Omega^2 R_{\alpha\beta}{}^{\gamma\delta} + 4(\Omega\nabla_{[\alpha} n^{[\gamma} - \frac{1}{2} n^2 \delta^{[\gamma}_{[\alpha})\delta^{\delta]}_{\beta]} \ . \tag{170}$$

On the other hand we may always write

$$R_{\alpha\beta}{}^{\gamma\delta} = C_{\alpha\beta}{}^{\gamma\delta} + 4P_{[\alpha}^{\ [\gamma}\delta^{\delta]}_{\beta]} \ , \tag{171}$$

where $C_{\alpha\beta\gamma\delta}$ is the Weyl tensor (the "traceless part" of the Riemann tensor), and similarly for the hatted curvature tensor. Hence we conclude that the Weyl tensor is a tensor of conformal weight minus two;

$$\hat{C}_{\alpha\beta\gamma}{}^\delta = C_{\alpha\beta\gamma}{}^\delta \ . \tag{172}$$

Moreover

$$\hat{P}_{\alpha\beta} = P_{\alpha\beta} + \frac{1}{\Omega}\nabla_\alpha\nabla_\beta\Omega - \frac{1}{2\Omega^2} g_{\alpha\beta} g^{\gamma\delta}\nabla_\gamma\Omega\nabla_\delta\Omega \ . \tag{173}$$

The tensor $P_{\alpha\beta}$ is simply related to the Ricci tensor, but the exact relation depends on the dimension D of spacetime. When $D = 4$ or 3 we have

$$P_{\alpha\beta} = \frac{1}{2}R_{\alpha\beta} - \frac{1}{12} g_{\alpha\beta} R \qquad\qquad D = 4 \tag{174}$$

$$P_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{4} g_{\alpha\beta} R \qquad\qquad D = 3 \ . \tag{175}$$

The significance of the Weyl tensor is that a spacetime of dimension $D > 3$ is conformally flat if and only if its Weyl tensor vanishes. In three dimensions the Weyl tensor vanishes identically, and its role is taken over by the conformally invariant Bach (or Cotton-York) tensor:

$$\hat{B}_{\alpha\beta\gamma} = B_{\alpha\beta\gamma} \equiv \nabla_{[\alpha}P_{\beta]\gamma} \qquad\qquad D = 3 \ . \tag{176}$$

A three dimensional spacetime is conformally flat if and only if the Bach tensor vanishes.

It is useful to record the once contracted Bianchi identity in four dimensions:

$$\nabla_\delta C_{\alpha\beta\gamma}{}^\delta + 2\nabla_{[\alpha}P_{\beta]\gamma} = 0 \qquad\qquad D = 4 \ . \tag{177}$$

In all dimensions the twice contracted Bianchi identity gives

$$\nabla_\beta P_\alpha{}^\beta = \nabla_\alpha P \ . \tag{178}$$

P is the trace of $P_{\alpha\beta}$.

Asymptotically anti-de Sitter spacetimes

We are now ready to discuss spacetimes that are asymptotically anti-de Sitter. Roughly speaking these should look like anti-de Sitter space far away from some central region. It is when this notion is to be made precise that the conformal viewpoint really pays off; as shown by Penrose the intuitive ideas can be captured in a few assumptions about the overall conformal structure. There is some latitude in deciding precisely what these assumptions are since we are trying to define a notion which is only approximate anyway—namely the notion of isolated gravitating systems. The assumptions are too weak if nothing can be proved, and they are too strong if no such systems occur in Nature. It is not *a priori* obvious that a middle ground exists, but it does.

At the outset we have a physical spacetime $\hat{M}$ with a metric $\hat{g}_{\alpha\beta}$ obeying Einstein's equations

$$\hat{R}_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}\hat{R} + \lambda\hat{g}_{\alpha\beta} = \hat{T}_{\alpha\beta} \ . \tag{179}$$

Next we identify $\hat{M}$ with the interior of a compact manifold-with-boundary M, whose boundary is $\mathcal{J}$. The manifold M has a metric $g_{\alpha\beta}$ which is conformally related to that of $\hat{M}$ by

$$g_{\alpha\beta} = \Omega^2\hat{g}_{\alpha\beta} \ . \tag{180}$$

It is assumed that

$$\Omega = 0 \quad \text{on} \quad \mathscr{J} \ . \tag{181}$$

The remarkable thing is that this assumption is all that is needed to ensure that $\hat{M}$ is asymptotically anti-de Sitter in a local sense, apart from necessary assumptions on regularity and analyticity as well as suitable fall-off assumptions for the energy-momentum tensor. Some extra assumptions will be needed to ensure that the topology is right—especially to ward of problems having to do with light rays that do not escape to $\mathscr{J}$—and more detailed assumptions concerning the existence of various conserved charges may be added as well. These complications should not be allowed to obscure the basic idea, which is that Einstein's vacuum equations for the physical metric are enough to ensure that the physical Weyl tensor must vanish in the neighbourhood of the hypersurface $\Omega = 0$ in the unphysical spacetime.

It is another question whether spacetimes with the required degree of regularity at $\mathscr{J}$ are in any sense generic as solutions of Einstein's equations. In some sense this is the question to which extent Einstein's equations themselves give rise to a suitable definition of isolated systems. Substantial progress has been made on this issue by Friedrich, with reassuring results at least in the anti-de Sitter case. But Friedrich's work is quite beyond our scope here—in fact we will make the necessary regularity assumptions without comments. In particular once it has been shown that Ω can be used as a coordinate in the neighbourhood of $\mathscr{J}$ I will tacitly assume that I can perform power series expansions in this coordinate, although as a matter of fact this is a dubious assumption.

The program is to deduce what Einstein's equations have to say about the unphysical geometry near $\mathscr{J}$. That is to say, our task is to remove the hats from various objects. We will make some simplifying assumptions—apart from analyticity assumptions the assumption that $\lambda < 0$ is the most important one, since the asymptotically flat case is significantly more subtle and requires further postulates, including topological assumptions and the assumption that the gradient of Ω is non-vanishing on $\mathscr{J}$. We also assume that $D = 4$ and that the energy-momentum tensor vanishes. The formulæ do get more cumbersome if matter is included, but in principle it is straightforward to formulate fall-off conditions that ensure that the main conclusions are unchanged.

Our starting point is therefore the pair of equations

$$\hat{R}_{\alpha\beta} = \lambda \hat{g}_{\alpha\beta} \quad \Leftrightarrow \quad \hat{P}_{\alpha\beta} = \frac{\lambda}{6}\hat{g}_{\alpha\beta} \tag{182}$$

and

$$P_{\alpha\beta} + \frac{1}{\Omega}\nabla_\alpha n_\beta - \frac{1}{2\Omega^2}n^2 g_{\alpha\beta} = \hat{P}_{\alpha\beta} \ . \tag{183}$$

We use the latter equation to transfer information to the unphysical geometry. We first combine our two equations and get the basic equation

$$\Omega^2 P_{\alpha\beta} + \Omega\nabla_\alpha n_\beta - \frac{1}{2}\Omega f g_{\alpha\beta} = 0 \ , \tag{184}$$

where

$$f = \frac{1}{\Omega}\left(n^2 + \frac{\lambda}{3}\right) \ . \tag{185}$$

One conclusion is immediate. If we let $\hat{=}$ stand for equality on $\mathcal{J}$ it must be true that

$$n_\alpha n^\alpha \hat{=} -\frac{\lambda}{3} \ . \tag{186}$$

Since we assume that $\lambda < 0$ the conclusion is that the surface $\Omega = 0$ must be timelike, just as in the case of anti-de Sitter spacetime itself. Note that the conclusion implies that Ω can be used as a coordinate in the neighbourhood of $\mathcal{J}$. It does not imply that the various geometrical objects that we encounter can be written in terms of power series expansions in Ω, but we will so assume.

We are now allowed to divide our basic equation by Ω. Contracting with the normal vector n_α we obtain

$$P_{\alpha\beta}n^\beta + \frac{1}{2}\nabla_\alpha f = 0 \ . \tag{187}$$

We can also take the curl of the basic equation. Then the Riemann tensor will arise from the term

$$\nabla_{[\alpha}\nabla_{\beta]}n_\gamma = \frac{1}{2}R_{\alpha\beta\gamma}{}^\delta n_\delta \ . \tag{188}$$

After a bit of massage the resulting equation is

$$\Omega \nabla_{[\alpha} P_{\beta]\gamma} + \frac{1}{2} C_{\alpha\beta\gamma}{}^{\delta} n_\delta = 0 \ . \tag{189}$$

A further conclusion follows:

$$C_{\alpha\beta\gamma}{}^{\delta} n_\delta \hat{=} 0 \quad \Rightarrow \quad C_{\alpha\beta\gamma}{}^{\delta} = \hat{C}_{\alpha\beta\gamma}{}^{\delta} = 0 \ . \tag{190}$$

The implication is easy given that a part of the Weyl tensor that is self dual in one pair of indices must be self dual in the other pair (but note that it holds only if the vector is spacelike or timelike). Since the Weyl tensor vanishes at infinity the spacetime is indeed asymptotically anti-de Sitter in a local sense.

Conserved charges

What are the structures available on $\mathscr{J}$? Are there conserved charges defined there ? The answer is that there are obvious candidates, and that they can be obtained in a few more steps. First we consider the intrinsic geometry on $\mathscr{J}$. In general, if we have a submanifold whose normal vector is n_α there is a projection operator that projects tensor fields to tensor fields on the submanifold:

$$q_\alpha^\beta = \delta_\alpha^\beta - \frac{1}{n^2} n_\alpha n^\beta \ . \tag{191}$$

The idea is that

$$T_{ab} \equiv q_a^\alpha q_b^\beta T_{\alpha\beta} \quad \Rightarrow \quad n^a T_{ab} = 0 \tag{192}$$

and so on; we use Latin indices as labels on tensors that have been projected down to the submanifold in this manner. Lowering one index on the projector we obtain the first fundamental form, that is to say that the intrinsic metric on $\mathscr{J}$ is

$$q_{ab} = g_{ab} + \frac{3}{\lambda} n_a n_b \ . \tag{193}$$

Since the normal vector is now spacelike the intrinsic metric is Lorentzian. Moreover the unphysical metric $g_{\alpha\beta}$ is defined only up to a conformal factor, so that it is only the conformal structure induced on $\mathscr{J}$ that is significant.

Next we observe that the vanishing of the Weyl tensor at $\mathcal{J}$ allows us to form the tensor

$$K_{\alpha\beta\gamma}{}^{\delta} \equiv \frac{1}{\Omega} C_{\alpha\beta\gamma}{}^{\delta} \ . \tag{194}$$

Using the equation that we have derived already together with the once contracted (unphysical) Bianchi identity we can show that

$$\nabla_\delta K_{\alpha\beta\gamma}{}^{\delta} = 0 \ . \tag{195}$$

Now consider the "electric" part of this rescaled Weyl tensor:

$$E_{ab} \equiv -\frac{3}{\lambda} K_{a\gamma b\delta} n^\gamma n^\delta \ . \tag{196}$$

Since this tensor is orthogonal to the normal it can be viewed as a symmetric and traceless tensor field on $\mathcal{J}$. Moreover we can introduce the unique torsion free and metric compatible derivative operator on $\mathcal{J}$, namely

$$D_a T_{b...}{}^{c...} = q_a^\alpha q_b^\beta ... q_\gamma^c ... \nabla_\alpha T_{\beta ...}{}^{\gamma ...} \ . \tag{197}$$

It is now a straightforward exercise to prove that

$$\nabla_\delta K_{\alpha\beta\gamma}{}^{\delta} = 0 \quad \Rightarrow \quad D_a E^{ab} = 0 \ . \tag{198}$$

By now we have a respectable amount of intrinsic structure defined on $\mathcal{J}$, including a metric, a metric compatible connection, and a transverse traceless tensor; it turns out to be enough to define conserved charges in some circumstances.

Suppose that the induced metric on $\mathcal{J}$ admits a conformal Killing vector ξ. Let C denote any spacelike surface on $\mathcal{J}$—a "cross section" of $\mathcal{J}$—and consider the integral

$$Q_\xi[C] = -\sqrt{\frac{-3}{\lambda}} \int_C E_{ab} \xi^a dS^b \ . \tag{199}$$

This is a conserved charge since we have, for two different cross sections bounding a volume V, that

$$Q_\xi[C] - Q_\xi[C'] = -\sqrt{\frac{-3}{\lambda}} \int dV D_a (E^{ab} \xi_b) = 0 \ . \tag{200}$$

The integrand is zero due to the facts that E^{ab} is transverse and traceless and that

$$D_{(a}\xi_{b)} = \alpha q_{ab} \ . \tag{201}$$

Since the conserved charges require a conformal Killing vector only, they will survive a conformal rescaling of the intrinsic metric on $\mathcal{J}$.

In the presence of matter the electric part of the rescaled Weyl tensor may or may not be divergence free, depending on the fall-off properties of the matter fields. On the other hand the magnetic part

$$B_{ab} \equiv -\frac{3}{\lambda} \star K_{a\gamma b\delta} n^{\gamma} n^{\delta} \tag{202}$$

of the rescaled Weyl tensor (where the star stands for the Hodge dual) is divergence free whether there is matter or no. This gives rise to an additional set of possible conserved charges, with no analogue in the asymptotically flat case (where they vanish identically).

Now there are two points worthy of note. First the result is unsatisfactory because there is no guarantee that the metric on $\mathcal{J}$ admits any conformal Killing vectors at all—we get conserved charges only in rather special cases. This suggests that the definition of asymptotically anti-de Sitter spacetimes ought to be strengthened by some further boundary condition, and we will presently turn to this question. The second point is that the result is surprisingly strong; our conserved charges are in fact analogous to the Bondi four-momentum in the asymptotically flat case, and the latter has the appealing property that the Bondi energy is a monotonically decreasing function of time. This reflects the fact that gravitational radiation can carry energy out of an asymptotically flat spacetime. As we have seen this does not happen in the asymptotically anti-de Sitter case, where the conclusion is that the charges are identically conserved in the absence of matter. The conclusion is that the attractive gravitational constant causes the gravitational radiation to fall back into the interior. In this sense the conformal boundary is much more like spatial infinity in the asymptotically flat case.

A stronger definition

We would like to strengthen the boundary conditions that define an asymptotically anti-de Sitter spacetime in such a way that the conserved charges always exist. This will be so if the geometry on $\mathscr{J}$ is conformally flat, that is to say if the three dimensional Bach tensor

$$B_{abc} \hat{=} 0 \ . \tag{203}$$

What is this condition when expressed in four dimensional language? The answer turns out to be

$$B_{ab} \hat{=} 0 \ . \tag{204}$$

Hence it is appropriate to strengthen the definition of asymptotically anti-de Sitter space by the requirement that the magnetic part of the rescaled Weyl tensor shall vanish on $\mathscr{J}$. In addition one may require that the topology of $\mathscr{J}$ shall be $\mathbf{R} \otimes \mathbf{S}^2$, although as we have seen the mathematics does not require this.

With this requirement added there is a certain universal structure in place, shared by all asymptotically anti-de Sitter spaces. This includes a set of ten conserved charges defined by means of the ten conformal Killing vectors that can be defined on a conformally flat 2+1 dimensional spacetime. They generate the conformal group $SO(3,2)$, and hence the asymptotic symmetry group is the anti-de Sitter group $SO(3,2)$. Again this is a much simpler result than that obtained in the asymptotically flat case where the group of asymptotic symmetries, i.e. the group leaving the universal structure invariant, is an infinite dimensional generalisation of the Poincaré group known as the BMS group. It is worthwhile to mention that the case of 2+1 dimensional anti-de Sitter spaces is special; in that case—which clearly requires a separate treatment—the conformal group acting on $\mathscr{J}$ is infinite dimensional, and so is the group of asymptotic symmetries acting on asymptotically anti-de Sitter spaces in 2+1 dimensions.

Finally, the canonical example of an asymptotically anti-de Sitter space is the Kottler (or Schwarzschild-anti de Sitter) solution

$$d\hat{s}^2 = -\left(1 - \frac{2m}{r} - \frac{\lambda}{3}r^2\right)dt^2 + \frac{dr^2}{\left(1 - \frac{2m}{r} - \frac{\lambda}{3}r^2\right)} + r^2 d\Omega^2 \ . \tag{205}$$

Let us define a new radial coordinate by

$$\rho = r^{-1} \tag{206}$$

and set

$$\Omega = \rho \ . \tag{207}$$

Then a quick calculation reveals that

$$d\hat{s}^2 \hat{=} \frac{\lambda}{3} dt^2 + \frac{d\rho^2}{\left(-\frac{\lambda}{3}\right)} + d\Omega^2 \ . \tag{208}$$

Evidently the metric induced on $\mathcal{J}$ is flat. Closer scrutiny reveals that there is a problem here: Like the Schwarzschild solution the Kottler solution has two asymptotic regions, so that the topology of the conformal boundary is not even connected. One reason why many definitions of asymptotically flat (or anti-de Sitter) spacetimes look a bit laboured is precisely that they are designed to handle such difficulties.

Afterthought

Throughout I have kept hinting at the fact that the analysis of infinity in asymptotically anti-de Sitter spaces is similar to but simpler than that of the asymptotically flat case. The negative cosmological constant serves as a regulator of the long distance behaviour of the gravitational field. In this respect it is of some interest to see what happens to Penrose's argument against the view that Einstein's theory can be regarded as an effective field theory of a spin 2 field defined on a flat background; this argument depends delicately on the long distance behaviour of the field and therefore it should be affected by a negative cosmological constant.

The original idea is as follows: Consider the Schwarzschild metric

$$d\hat{s}^2 = -\left(1 - \frac{2m}{r}\right) dt^2 + \frac{dr^2}{\left(1 - \frac{2m}{r}\right)} + r^2 d\Omega^2 \ . \tag{209}$$

This may be glued to an interior solution describing the sun, so no fancy topology is being assumed. Now suppose that there is an underlying flat metric, and that whatever theory that gives rise to the Schwarzschild metric as an "effective" metric respects causality as defined by the flat metric—we intend to think of gravity much as we think of electrodynamics in a medium, where the velocity of light can differ from that *in vacuo*, but cannot exceed it. It follows that a vector that is timelike or lightlike with respect to the Schwarzschild metric has to be timelike or lightlike with respect to the flat background metric as well. This appears to be so if we choose the background metric to be

$$d\hat{\sigma}^2 = -dt^2 + dr^2 + r^2 d\Omega^2 \ . \tag{210}$$

There is however a catch. Consider an outgoing light ray in the Schwarzschild geometry, obeying the equation

$$\frac{dt}{dr} = \frac{r}{r - 2m} \ . \tag{211}$$

The solution is

$$u = t - r - 2m \ln{(r - 2m)} \ , \tag{212}$$

where u is a constant. We refer to it as the retarded time of the light ray; it is a useful coordinate on $\mathscr{J}^+$, future null infinity of the Schwarzschild spacetime. But we can also introduce a future null infinity by means of our flat background metric. The retarded time of an outgoing null geodesic with respect to the background metric is

$$u_0 = t - r \ , \tag{213}$$

and clearly

$$\lim_{r \to \infty} (u_0 - u) = \infty \ . \tag{214}$$

The conclusion is that the light ray does not hit future null infinity—as defined by the background metric—at all, rather it ends up at future temporal infinity. But it does seem reasonable to demand that the background spacetime should agree with the "effective" physical spacetime at infinity, where the effects of the gravitating mass are negligible. The background metric that we selected fails to do this.

So we try again. We introduce a new radial coordinate

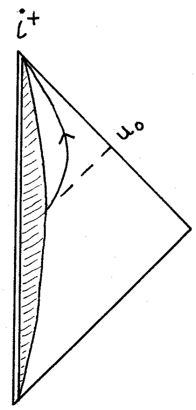

Figure 10: A star in a flat background. The trouble is that light from the star ends up at i^+.

$$r' = r + 2m \ln (r - 2m) \ . \tag{215}$$

As our flat background metric, we choose

$$d\hat{\sigma}'^2 = -dt^2 + dr'^2 + r'^2 d\Omega^2 \ . \tag{216}$$

This resolves our difficulty, since with this choice of flat metric the retarded times of the radial light ray becomes the same with respect to both metrics. The metrics have been properly connected at infinity. But now another problem hits us. If $d\hat{s}^\alpha$ is some tangent vector which is lightlike with respect to the Schwarzschild metric, we must have that

$$d\hat{s}^2 = (1 - \frac{2m}{r})(-dt^2 + dr'^2) + r^2 d\Omega^2 = 0 \ . \tag{217}$$

But with respect to the our new background metric we then find that

$$d\hat{\sigma}'^2 = \left((r + 2m \ln (r - 2m))^2 - \frac{r^3}{r - 2m} \right) d\Omega^2 \ . \tag{218}$$

For large r this goes like

$$\frac{2mr^2}{r - 2m}(2 \ln r - 1) > 0 \ . \tag{219}$$

Hence geodesics which are null but not radially directed with respect to the Schwarzschild metric are spacelike with respect to our new proposed background metric. But we already argued that the "true" background metric cannot have this property, and therefore our second candidate also fails.

The failure is not due to lack of imagination on our part, because Penrose went on to prove that it is impossible to define a flat metric on the Schwarzschild spacetime which fulfils both our requirements: It should lead to the same notion of null infinity as the Schwarzschild metric, and a curve which is causal with respect to the latter should also be causal with respect to the former. Hence we have a definite argument against viewing Einstein's theory as a kind of effective theory of a spin 2 field defined with respect to an unobservable flat metric. Such a flat background metric simply cannot exist.

What happens if the cosmological constant is negative? In that case the behaviour far from the sun will be dominated by the cosmological constant, rather than by the sun itself. But this suggests that the sun should be more or less irrelevant for the causal structure far from the sun, and that Penrose's argument should fail in this case. Indeed this is so. We replace the Schwarzschild metric with the Kottler metric, and choose an anti-de Sitter background metric by setting $m = 0$ in the Kottler metric. Repeating the calculation that was done above leads—once the appropriate integral has been performed—to

$$u_0 - u \sim \frac{1}{r} \tag{220}$$

for large values of r.[2] This means that $\mathscr{J}$ defined with respect to one of the metrics agrees with $\mathscr{J}$ defined with respect to the other, and that Penrose's argument indeed fails in this case.

The lesson is again that infinity in anti-de Sitter space is very far away—it is so far away that gravitational radiation cannot reach it, and the causal structure in its vicinity is unaffected by any mass concentration in its interior.

[2]This is a bit confused. For a correct version of the argument, see K. Öberg: *Backgrounds for the Schwarzschild solutions; the influence of* λ, Master's Thesis, Stockholm/Lund 2000.

Exercise:

- Prove that

$$\nabla_{[\alpha} R_{\beta\gamma]\mu\nu} = 0 \quad \Rightarrow \quad \nabla_\delta C_{\alpha\beta\gamma}{}^\delta + 2\nabla_{[\alpha} P_{\beta]\gamma} = 0 \ . \tag{221}$$

- Given a hypersurface defined by $\Omega = 0$ and with normal vector $n_\alpha = \nabla_\alpha \Omega$. Show that D_a as defined in the text is the standard metric compatible covariant derivative formed from the induced metric q_{ab}. (Hint: You are supposed to know that this derivative exists, and that it is unique.) Also check that

$$\nabla_\alpha K^\alpha{}_{\beta\gamma\delta} = 0 \quad \Rightarrow \quad D_a E^{ab} = 0 \ , \tag{222}$$

where the tensors and the derivative operators are defined in the text.

GREEN FUNCTIONS

Generalities

No discussion of the geometry of anti-de Sitter space would be complete without some mention of how the wave equation behaves on such a background—we want to know what the geometry does, not only what it is. It will prove convenient to begin with a discussion of the Laplace equation on hyperbolic space, partly since this is of interest in itself and partly because we can then approach the wave equation by means of an analytic continuation from hyperbolic space.

In a general curved space the invariant Laplace operator is defined as

$$\triangle \equiv D_a D^a = \frac{1}{\sqrt{g}} \partial_a (\sqrt{g} g^{ab} \partial_b) \tag{223}$$

and the equation that we wish to solve is

$$(\triangle - \mu^2)\phi = 0 \ . \tag{224}$$

We will refer to this as the Helmholtz equation. In flat space we obtain the Klein-Gordon equation after an analytic continuation to Minkowski space, and this is the reason why we have chosen a notation that suggests that the eigenvalue μ^2 is greater than zero, but this is not necessarily so. In flat space we know that $\mu^2 > 0$ leads to solutions that fall off exponentially at infinity, while $\mu^2 < 0$ leads to oscillatory solutions. There is a similar division into two main cases in hyperbolic space, but the "critical" value is negative and there is an interesting "fine structure" just above it.

The discussion will focus on the Green function. By definition it obeys

$$(\triangle - \mu^2) G(1,2) = \frac{1}{\sqrt{g}} \delta(1,2) \ . \tag{225}$$

The equation is to be solved under the condition that the Green function vanishes when one of its arguments lies on some closed hypersurface, or else its normal derivative is to vanish there. To see why the Green function is a good thing to have, indeed the only thing one needs, suppose that A is such a closed hypersurface surrounding a volume V. Then we can explicitly construct a field ϕ that obeys the homogeneous equation, using only data on the boundary A. If the value of the field is specified on the boundary this is called Dirichlet data, if its normal derivative is specified it is called Neumann data. The construction, given a Green function with the stated properties, is as follows: Let the second argument x of the Green function be some point inside the volume and let ϕ be any solution of the homogeneous equation. Then

$$\begin{aligned} \phi(x) &= \int dV (\phi(\triangle - \mu^2)G - (\triangle - \mu^2)\phi G) = \\ &= \int dV D^a(\phi D_a G - D_a\phi G) = \int dA^a(\phi D_a G - D_a \phi G) \ . \end{aligned} \tag{226}$$

If $G = 0$ on the boundary this gives the field ϕ at x in terms of Dirichlet data, and if the normal derivative of G vanishes on the boundary we get ϕ in terms of Neumann data. More general "mixed" boundary conditions can also be contemplated.

Given Dirichlet data on the boundary our solution is unique if $\mu^2 > 0$. This follows from another simple observation, viz.

$$\int dV (D_a\phi D^a\phi + \mu^2\phi^2) = \int dV D^a(\phi D_a\phi) = \int dA^a \phi D_a \phi \ . \tag{227}$$

All the terms in the integrand on the left hand side are positive or zero, while the right hand side is zero. It follows that the only solution that vanishes on the boundary has vanishing gradient everywhere, therefore the solution is zero everywhere. If Neumann data are specified on the boundary the solution is unique up to a constant. If $\mu^2 < 0$ no conclusion can be drawn.

In our discussion we will place the enclosing hypersurface at infinity and the Green function will be assumed to vanish at infinity. This is not quite the situation that we have just considered, but it will turn out that for a special value of μ^2 (the "conformally coupled scalar") the discussion is directly relevant because we can transform the whole problem to that of solving the Helmholtz equation on the conformally compactified space, which has a boundary at finite distance from any point in the interior.

Flat space

For reference, let us collect some relevant results in flat space. The question is to which extent these results can be generalized. The first striking thing about flat space is that is an isotropic space—all directions are equivalent—and this means that the Green function can depend only on the geodesic distance between the source and observation points. It follows that the partial differential equation that we have to solve can be reduced to an ordinary differential equation; if we choose spherical polars with the origin at the source point it will be enough to solve

$$\left(\partial_r^2 + \frac{n-1}{r}\partial_r - m^2\right) G_n(r;m^2) = \delta(r) \ . \tag{228}$$

(We use m^2 rather then μ^2 for the eigenvalue, since m can be identified with mass in a fairly unproblematic way in flat space.) Since hyperbolic space is isotropic as well, this feature generalizes. If you like you can think of the Green function as a solution of the homogeneous equation in the interval $0 < r < \infty$. We require that it falls to zero at infinity and diverges in a suitable way at the origin.

The behaviour at $r = 0$ can be gleaned by solving the case $m^2 = 0$, and this happens to be trivial. If $n = 2$ it is

$$G(r) = \frac{1}{2\pi} \ln r + \text{constant} \ . \tag{229}$$

The catch is that this cannot be reconciled with the requirement that the Green function is to fall to zero at infinity. In two dimensions the Green function of the Laplace equation has an incurable infrared divergence (although it can be defined in a box). This difficulty will actually go away in hyperbolic space. If $n > 2$ the solution is

$$G(r) = -\frac{1}{(n-2)A_{n-1}} \frac{1}{r^{n-2}} + \text{constant} \ , \tag{230}$$

where A_n is the area of the n-sphere. With a little effort one can show that

$$A_n = 2\frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})} \ . \tag{231}$$

If we set the arbitrary constant to zero we are home.

For non-vanishing m^2 the solution is not elementary. The answer (now vanishing at infinity in all cases) turns out to be

$$G_n(r;m^2) = -\frac{i}{4}\left(\frac{im}{2\pi}\right)^{\nu}\frac{1}{r^{\nu}}H^{(1)}_{\nu}(imr) \ , \qquad \nu = \frac{n-2}{2} \ , \tag{232}$$

where $H^{(1)}$ denotes a Hankel function of the first kind, with the asymptotic behaviour

$$H^{(1)}_{\nu}(z) \sim \sqrt{\frac{2}{\pi z}}e^{i(z-(\nu+\frac{1}{2})\frac{\pi}{2})} \ . \tag{233}$$

As advertized our Green functions vanish exponentially for large r provided that m is real; if $m^2 < 0$ we obtain an oscillatory behaviour.

We have insisted on keeping the dimension n arbitrary. Examination of the equation reveals that the solution in $n+2$ dimension follows from that in n dimensions if we set

$$G_{n+2}(r;m^2) = -\frac{1}{2\pi r}\partial_r G_n(r;m^2) \ . \tag{234}$$

(The factor in front is chosen so as to obtain the correct strength of the singularity.) Hence all the Green functions can be derived from G_2 and G_3 by repeated differentiation. This statement is consistent with the explicit solution because of recurrence relations obeyed by all the Bessel functions; it is also a feature that generalizes to hyperbolic space.

In odd dimensions things simplify because the asymptotic expansion of a Hankel function of half integer order terminates after a finite number of terms. Thus we get

$$G_3(r;m^2) = -\frac{1}{4\pi}\frac{e^{-mr}}{r} \ . \tag{235}$$

A similar simplification occurs in odd dimensional hyperbolic space, where the Green functions can again be expressed in terms of elementary functions.

If we have a solution of the Laplace equation with the same fall-off behaviour as the Green function we would conclude that it is square integrable if $m > 0$ but not if $m = 0$. On the other hand the energy integral is convergent also for $m = 0$.

The equation in $\mathbf{H}^n$

Like flat space, hyperbolic space is isotropic—therefore the Green function can only depend on the geodesic distance d between the source point and the observation point, and the calculation can again be reduced to that of finding the Green function of an ordinary differential equation. There are two natural ways to arrive at this. We can choose a geodesic polar coordinate system with the origin at the source point, so that the Green function is a function of the coordinate r only. The equation to be solved is

$$(\partial_r^2 + (n-1)\coth r\partial_r - \mu^2)G(r;\mu^2) = \delta(r) \ . \tag{236}$$

Since we are using coordinates where $r = d$, we get the Green function in coordinate independent form by replacing r with d.

Alternatively we can use the embedding coordinates. We first define a projection operator

$$q_\alpha{}^\beta = \delta_\alpha^\beta + X_\alpha X^\beta \ , \qquad q_\alpha{}^\alpha = n \ . \tag{237}$$

This projection operator has the property that it annihilates the normal vector of the quadric hypersurface that we are interested in ($\mathbf{H}^n$ for the present, although the formulæ work for $\mathbf{adS}_n$ as well). That is to say that

$$q_\alpha{}^\beta X_\beta = 0 \ . \tag{238}$$

The standard covariant derivative on the hypersurface is now given by

$$D_a = q_a{}^\beta \partial_\beta \ . \tag{239}$$

It is a straightforward exercise to show that

$$\triangle = q^{\alpha\beta}\partial_\alpha\partial_\beta + nX^\alpha\partial_\alpha = -\frac{1}{2}J_{\alpha\beta}J^{\alpha\beta} \ . \tag{240}$$

The second equality makes contact with group theory since J^2 is a Casimir operator of the isometry group; this makes sense since the Laplacian in flat space is

$$\triangle = -P^2 \ , \tag{241}$$

where P^2 is a Casimir operator of the Euclidean group. A great deal can be learned about our problem from a study of the representation theory of the relevant group, but we will not rely on this here.

The Green function is assumed to be a function of geodesic distance d only. It is convenient to use the variable

$$u = \cosh d = -X_1 \cdot X_2 \ . \tag{242}$$

It follows that when $u \neq 1$

$$(\triangle - \mu^2)G(u;\mu) = -(1-u^2)G''(u) + nuG'(u) - \mu^2 G(u) = 0 \ . \tag{243}$$

This is precisely what the equation that we had above turns into, if we make the substitution $u = \cosh r$. When $n = 2$ it is recognizable as Legendre's equation. The solution is to be sought under the conditions that

$$\lim_{u\to\infty} G(u) = 0 \qquad\qquad \lim_{u\to 1} G(u) = -\frac{1}{(n-2)A_{n-1}}\frac{1}{r^{n-2}} \ , \tag{244}$$

unless $n = 2$—the general rule is that the strength of the singularity is the same as in flat space, since any curved space is locally flat.

Of course our differential equation can be transformed this way and that through various substitutions. One frequently sees

$$v = -\sinh^2\frac{r}{2} \ \Rightarrow \ -v(1-v)G''(v) - \frac{n}{2}(1-2v)G'(v) - \mu^2 G(v) = 0 \ . \tag{245}$$

This is recognizable as the hypergeometric equation for a special choice of parameters. We prefer the substitution $u = \cosh r$ however.

The solution in $\mathbf{H}^n$

As in flat space the two and three dimensional Green functions are the keys to the general solution because the higher dimensional cases can be obtained from G_2 or G_3 by repeated differentiation. In the two dimensional case the defining differential equation is precisely Legendre's equation

$$(1-u^2)Q_\nu''(u) - 2uQ_\nu'(u) + \nu(\nu+1)Q_\nu(u) = 0 \ ; \tag{246}$$

all that we have to do is to adjust ν in terms of μ^2. Because we require the solution to be singular at $u = 1$ the solution that we want is a Legendre function of the second kind, called Q_ν. Indeed

$$G_2(r;\mu^2) = -\frac{1}{2\pi} Q_\nu(\cosh r) \ . \tag{247}$$

Here we must use

$$\nu(\nu+1) = \mu^2 \quad \Leftrightarrow \quad \nu = -\frac{1}{2} + \sqrt{\mu^2 + \frac{1}{4}} \ . \tag{248}$$

Since I do not wish to go into any discussion of special functions—and since the three dimensional case can be solved with elementary functions—I will not discuss this solution for general values of μ^2.

It is good to understand the case $\mu^2 = 0$ though, the more so because this case will found to be of special interest later. An alternative way to get the answer is then to use Gauss' theorem and place the boundary of the volume at constant r;

$$1 = \int dV \triangle G = \int dA^a \sqrt{g} D_a G = 2\pi \sinh r \partial_r G_2(r) \ . \tag{249}$$

In this way the geometry requires that

$$G_2(r;0) = \frac{1}{2\pi} \int^r \frac{dr}{\sinh r} \ . \tag{250}$$

The integral is elementary and the answer is

$$G_2(r;0) = \frac{1}{2\pi} \ln \tanh \frac{r}{2} = -\frac{1}{4\pi} \ln \frac{\cosh r + 1}{\cosh r - 1} = -\frac{1}{2\pi} Q_0(\cosh r) \ , \tag{251}$$

where Q_0 is a Legendre function of the second kind, as advertized. Unlike in flat space there are no infrared problems to worry about here.

The general solution when $n = 3$ can be found by a little experimentation, starting from the $\mu^2 = 0$ case which can be worked out using Gauss' theorem as above. The answer can be expressed in terms of elementary functions, as advertized:

$$G_3(r;\mu^2) = -\frac{1}{4\pi} \frac{e^{\pm r\sqrt{\mu^2+1}}}{\sinh r} \ . \tag{252}$$

This is clearly reminiscent of the Green functions in flat space, but there are some noteworthy differences too. The "critical" value below which the oscillatory behaviour appears is $\mu^2 = -1$ rather than zero (as in flat space); moreover in the range

$$-1 \leq \mu^2 < 0 \tag{253}$$

both branches of the square root are consistent with exponential fall-off to zero at infinity. On the other hand if we consider a function with this fall-off behaviour it will be square integrable on $\mathbf{H}^3$ only if we choose the negative sign in the exponent. We will have more to say about where this non-uniqueness comes from later on. Meanwhile it is interesting to confirm that oscillatory solutions appear below $\mu^2 = -1$. If you remember the discussion of plane waves on the Poincaré disk you see that we can do so by choosing half space coordinates and making the Ansatz

$$\phi = x^s \ . \tag{254}$$

The solution is assumed to be constant on the horospheres, and we must have a complex s in order to get oscillatory solutions. The Helmholtz equation is

$$\triangle\phi = \mu^2\phi \qquad \Rightarrow \qquad x^2(\partial_x^2 + \partial_y^2 + \partial_z^2)\phi - x\partial_x\phi = \mu^2\phi \tag{255}$$

and this leads to

$$s(s-2) = \mu^2 \qquad \Leftrightarrow \qquad s = 1 \pm \sqrt{1+\mu^2} \ . \tag{256}$$

Our point has been proved.

The higher dimensional cases can again be found by repeated differentiation of G_2 and G_3 if a minor complication is kept in mind; it is easy to show that

$$G_{n+2}(u;\mu^2) = -\frac{1}{2\pi}\partial_u G_n(u;\mu'^2) \ , \tag{257}$$

where the strength of the singularity has been properly adjusted as well. The complication is that μ^2 on the left hand side is not the same as μ'^2 on the right hand side. Taking this into account we find the answer for an arbitrary even dimension $n = 2p$:

$$G_{2p}(r;\mu^2) = -\frac{1}{2\pi}\left(-\frac{1}{2\pi\sinh r}\partial_r\right)^{p-1} Q_\nu(\cosh r) \ , \tag{258}$$

where

$$\nu = -\frac{1}{2} + \sqrt{\mu^2 + \frac{(n-1)^2}{4}} \ . \tag{259}$$

And we also find the answer for an arbitrary odd dimension $n = 2p+1$:

$$G_{2p+1}(r;\mu^2) = -\frac{1}{4\pi}\left(-\frac{1}{2\pi \sinh r}\partial_r\right)^{p-1}\left(\frac{e^{-r\sqrt{\mu^2+\frac{(n-1)^2}{4}}}}{\sinh r}\right) \ . \tag{260}$$

The remarks about the range of μ^2 and the two branches of the square root can evidently be repeated.

Conformally coupled scalars

The discussion so far has revealed that it is the behaviour at infinity that has to be understood, and this is the kind of question that can be clarified by going to the conformally compactified space. In general a solution to the Helmholtz equation on the physical space will not correspond to a solution of the Helmholtz equation on the compactified space, but there is one exception that is called the conformally coupled scalar. Consider the following natural equation:

$$(\triangle - kR)\phi = 0 \ . \tag{261}$$

Here R is the curvature scalar and k is some constant. Whatever the value of k this will be the Laplace equation in flat space; it can be derived from the action

$$S = \int \sqrt{g}(D_a\phi D^a\phi + k\phi^2 R) \ . \tag{262}$$

It can be shown that this action is invariant under the conformal rescalings

$$\hat{g}_{ab} = \Omega^2 g_{ab} \qquad \hat{\phi} = \Omega^w \phi \tag{263}$$

(where w is called the conformal weight of the scalar field) if and only if

$$k = \frac{n-2}{4(n-1)} \qquad \text{and} \qquad w = \frac{2-n}{2} \ . \tag{264}$$

We adopt these values of k and w from now on; they define the conformally coupled scalar. It follows that $\hat{\phi}$ will be a solution of our equation on the conformally related space, given that ϕ is a solution on the space we started out from. The calculation needed for the proof is simple, given that (on a space of dimension n)

$$\hat{R}(\hat{g}) = \frac{1}{\Omega^2}R(g) + \frac{2}{\Omega^3}(1-n)\nabla_a\nabla^a\Omega + \frac{1}{\Omega^4}(1-n)(n-4)\nabla_a\Omega\nabla^a\Omega \ , \tag{265}$$

where the original metric is used to raise and lower indices on the right hand side.

Let us now specialize to hyperbolic space. With the normalisation that we have adopted the Riemann curvature scalar of $\mathbf{H}^n$ is

$$R = -n(n-1) \ . \tag{266}$$

Hence the equation for a conformally coupled scalar on $\mathbf{H}^n$ is

$$\left(\triangle + \frac{n(n-2)}{4}\right)\phi = 0 \ . \tag{267}$$

Note that this implies that the conformally coupled scalar has

$$\mu^2 = -\frac{n(n-2)}{4} \qquad \Rightarrow \qquad \mu^2 + \frac{(n-1)^2}{4} = \frac{1}{4} \ . \tag{268}$$

For reference, $\mu^2 = 0$ if $n = 2$, $\mu^2 = -3/4$ if $n = 3$ and $\mu^2 = -2$ if $n = 4$.

Finally we observe that a Green function is a solution of the homogeneous equation except at the source and observation points. For a conformally coupled scalar it must therefore be true that

$$\hat{G}(1,2) = \Omega^w(1)G(1,2)\Omega^w(2) \ , \tag{269}$$

where $\hat{G}$ is the Green function in a space that is conformally related to another space where the Green function is G. In particular this means that for the special value of μ^2 given above the Green function in hyperbolic space can be computed in terms of the massless Green function in flat space, which is nice.

The conformally coupled scalar on $\mathbf{H}^n$

The Green function G of the conformally coupled scalar on $\mathbf{H}^n$ can evidently be derived from the Green function $G^{\mathbf{E}}$ of the Laplacian on the conformally related space $\mathbf{E}^n$. The metric on $\mathbf{H}^n$ is in stereographic coordinates

$$g_{ab} = \Omega^2 \delta_{ab} = \left(\frac{2}{1-\rho^2} \right)^2 \delta_{ab} \ . \tag{270}$$

What we must keep in mind is that we are no longer interested in flat space Green functions that vanish at infinity; infinity in $\mathbf{H}^n$ now corresponds to the hypersurface $\rho = 1$, that is to say to the surface of the unit sphere in $\mathbf{E}^n$.

For the comparison with our previous results we just have to remember that the radial coordinate ρ is related to the geodesic distance d (that is r) from the origin by

$$\rho = \tanh \frac{d}{2} = \tanh \frac{r}{2} \ . \tag{271}$$

If we want to express the answer in embedding coordinates we use

$$\rho^2 = \frac{\cosh r - 1}{\cosh r + 1} = \frac{X_1 \cdot X_2 + 1}{X_1 \cdot X_2 - 1} = -\frac{(X_1 - X_2)^2}{(X_1 + X_2)^2} \ . \tag{272}$$

We can now solve for the Green function in our favourite dimensions. In two dimensions the conformal weight of the conformally coupled scalar is zero, so the Green functions on two conformally related spaces are in fact equal—no rescaling is needed. If we require the flat space Green function to vanish at the boundary (so that the solutions are determined by Dirichlet conditions on the conformal boundary) we obtain in $\mathbf{H}^2$

$$G_2(1,2) = G_2^{\mathbf{E}}(1,2) = \frac{1}{2\pi} \ln \rho + \text{constant} = \frac{1}{2\pi} \ln \tanh \frac{r}{2} \ , \tag{273}$$

where the constant was adjusted so that

$$G_2^{\mathbf{E}}(\rho = 1) = 0 \ . \tag{274}$$

To solve for the Green function in $\mathbf{H}^2$ is the same problem as that of solving for the Green function in a box in flat space, that is why no infrared problems occur.

When $n > 2$ the flat space Green function that we need for Dirichlet boundary conditions is

$$G_2^{\mathbf{E}}(\rho = 1) = 0 \quad \Rightarrow \quad G_n^{\mathbf{E}} = -\frac{1}{(n-2)A_{n-1}}\left(\frac{1}{\rho^{n-2}} - 1\right) \ , \tag{275}$$

where A_{n-1} is the area of $\mathbf{S}^{n-1}$. The conformal weight of the scalar is $w = (2-n)/2$, so the Green function in hyperbolic space is

$$G_n(\rho) = \left(\frac{2}{1-\rho^2}\right)^{\frac{n-2}{2}} G_n^{\mathbf{E}}(\rho)\ 2^{\frac{n-2}{2}} = -\frac{2^{2-n}}{(n-2)A_{n-1}}(1-\rho^2)^{\frac{n-2}{2}}\frac{1-\rho^{n-2}}{\rho^{n-2}} \ . \tag{276}$$

It can be shown that this agrees with our previous solution when the appropriate conformal value of μ^2 is inserted in the latter. The solution is unique so as yet there is no trace of the mysterious "second branch" of non-square integrable solutions.

Let us specialize to $n = 3$ to simplify matters a bit. Now we can, if we like, contemplate more general boundary conditions on G_3 in flat space, such as

$$G_3^{\mathbf{E}} = -\frac{1}{4\pi}\left(\frac{1}{\rho} + k\right) \quad \Rightarrow \quad G_3^{\mathbf{E}}(\rho = 1) = \text{constant} \ . \tag{277}$$

We get a Green function for all values of the constant k, although it vanishes at the boundary only if $k = -1$. When expressed in terms of r the Green function becomes

$$G_3(r) = -\frac{1+k}{8\pi}\frac{e^{\frac{r}{2}}}{\sinh r} - \frac{1-k}{8\pi}\frac{e^{-\frac{r}{2}}}{\sinh r} \ . \tag{278}$$

This is a linear combination of the two solutions that we found earlier. From our discussion of the boundary value problem in general we see that if the

Green function does not vanish on the boundary—and its normal derivative neither—then we are dealing with a mixed boundary value problem in the conformally compactified space. The uniqueness of the Dirichlet problem is gone, but otherwise there is nothing wrong with it—and our argument for uniqueness is gone anyway since $\mu^2 < 0$.

Having understood why the Dirichlet case $k = -1$ is not unique, we return to it and contemplate the Green function in embedding coordinates. A straightforward calculation shows that

$$G_3(1,2) = -\frac{1}{4\pi}\left(\frac{1}{\sqrt{(X_1 - X_2)^2}} - \frac{1}{\sqrt{-(X_1 + X_2)^2}}\right) . \tag{279}$$

This expression highlights an interesting fact: The second term here is regular everywhere on $\mathbf{H}^3$ and it is in fact a solution to the homogeneous equation (as one readily checks); hence it can be added to the Green function with an arbitrary coefficent. It is singular when $X_1^\alpha = -X_2^\alpha$, that is if (say) the observation point is in $\mathbf{H}^3$ and the source point is on the other, discarded, sheet of the hyperboloid in embedding space. It can be thought of as an "image charge" that is added outside our space to ensure the correct behaviour at the boundary. In fact this interpretation is exactly right. By means of a few manipulations we can bring the Green function (for the Dirichlet case say, and with the source and observation points at arbitrary position) to the form

$$G(1,2) = 2^{-1}\sqrt{1-\rho_1^2}\,G_3^{\mathbf{E}}(1,2)\sqrt{1-\rho_2^2} \ , \tag{280}$$

where the Green function in Euclidean space is

$$G_3^{\mathbf{E}}(1,2) = -\frac{1}{4\pi}\left(\frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} - \frac{1}{|\rho_2\mathbf{x}_1 - \frac{1}{\rho_2}\mathbf{x}_2|}\right) . \tag{281}$$

This is the Green function for the Laplacian in flat space including an image charge that ensures that the Green function vanishes on the unit sphere—the exterior of the unit sphere is conformally related to the discarded sheet of the hyperboloid, so it checks.

The story is similar if $n \neq 3$. For reference we give the (particularly appealing and simple) Green function for the conformally coupled scalar with

Dirichlet conditions imposed on the boundary of the conformally compactified space when $n = 4$, expressed first in terms of the variable $u = \cosh r$ and then in embedding coordinates:

$$G_4 = -\frac{1}{8\pi^2}\left(\frac{1}{u-1} + \frac{1}{u+1}\right) = -\frac{1}{4\pi^2}\left(\frac{1}{(X_1 - X_2)^2} + \frac{1}{(X_1 + X_2)^2}\right). \tag{282}$$

And our understanding of Green functions on $\mathbf{H}^n$ is as complete as we care to make it.

Flat spacetime

The main difference between the Euclidean and the Lorentzian case is that there is a whole zoo of Green functions in the latter. This comes about because there are homogeneous solutions of a hyperbolic equation, and given a Green function we obtain another by adding a solution to the homogeneous equation. On the other hand the only solution of the homogenous Laplace equation (given that we have imposed suitable boundary conditions at infinity) is zero, which means that the Euclidean Green function is unique. Now we are interested in the situation where the Lorentzian space can be reached from a space with positive definite metric through analytic continuation. The analytic continuation will then give rise to a Lorentzian Green function that is distinguished in the sense that it arises as the analytic continuation of the Euclidean Green function. To be precise about it, suppose that

$$t = -it_E \tag{283}$$

where t_E is the Euclidean "time", and let G_E be the Green function of the Laplacian. Then the Feynman propagator

$$G^F(\mathbf{x}, t) = iG^E(\mathbf{x}, it) \tag{284}$$

is a Green function of the wave equation. By construction it is analytic in the second and fourth quadrant of the complex t-plane. Other Green functions have other analyticity properties, for instance the retarded Green function can be defined as that Green function which is analytic in the lower complex t-plane.

Since Minkowski space is isotropic the Green functions are functions of

$$\sigma = g_{\alpha\beta}x^\alpha x^\beta \tag{285}$$

only (where we assume that the origin of the coordinate system has been place at the source point). What is new compared to the Euclidean case is that σ can take negative values. Moreover it will vanish whenever the separation between the source and observation points is lightlike, and this means that the Green function will be singular there. The presence of these singularities resolves a puzzle: Since the Euclidean Green function is real it would seem that the Feynman propagator must be imaginary, and then it is difficult to see how it can be the Green function of a real equation. The resolution is that the combination of its analyticity properties and its singularities will force it to develop a real part. (The meaning of this mystical statement will be clear in the examples.)

The Feynman propagator is by construction analytic in the second and fourth quadrant of the complex t-plane. Moving the singularities off the real axis by adding an infinitesimal imaginary part to t we find

$$\mathbf{x}^2 - (t - \mathrm{sgn}(t)i0)^2 = \mathbf{x}^2 - t^2 + i0 = \sigma + i0 \ . \tag{286}$$

(The "$i0$" notation keeps track of the analyticity properties.) Hence we can just as well say that the Feynman propagator is defined as that Green function that is analytic in the upper complex σ-plane.

Let us now split the Feynman propagator into real and imaginary parts:

$$G^F = \bar{G} + \frac{i}{2}G^{(1)} \ . \tag{287}$$

The parts are called the symmetric Green function and Hadamard's elementary function, respectively. The latter is a solution of the homogeneous wave equation. Other well known Green functions can be derived by multiplying with suitable step functions. Thus the retarded Green function is

$$G^R = 2\theta(t)\bar{G} \ , \tag{288}$$

and the commutator Green function is

$$\tilde{G} = 2\epsilon(t)\bar{G} \ . \tag{289}$$

Recall that—despite its quantum field theoretical name—the properties of the commutator Green function is of crucial importance for causal propagation in classical field theory as well. Consider the Cauchy problem

$$\phi(\mathbf{x},0) = u(\mathbf{x}) \qquad\qquad \partial_t\phi(\mathbf{x},t) = v(\mathbf{x}) \tag{290}$$

(initial data chosen on the spatial hypersurface $t = 0$, say). Then there is a unique solution of the Klein-Gordon equation given by

$$\phi(\mathbf{x},t) = \int_{t'=0} d^dx'(\partial_{t'}\tilde{G}(x,x')u(\mathbf{x}') - \tilde{G}(x,x')v(\mathbf{x}')) \ . \tag{291}$$

Physical effects will therefore propagate in a causal manner if and only if the commutator Green function vanishes outside the lightcone. We will see that this is true for the Klein-Gordon equation; there is the added benefit that the step functions do not disturb the Lorentz invariance of these Green functions.

The simplest illustration of the general discussion is provided by the wave equation in four dimensions. Place the origin at the source point, and define

$$\sigma = g_{\alpha\beta}x^\alpha x^\beta \ . \tag{292}$$

Then

$$G_4^F = -\frac{i}{4\pi^2}\frac{1}{\sigma + i0} = -\frac{1}{4\pi}\delta(\sigma) - \frac{i}{4\pi^2}\frac{1}{\sigma} \ . \tag{293}$$

The singularities are poles, and the support of $\bar{G}$—and hence of the commutator and retarded Green functions—is not only vanishing outside the light cone, it is actually confined to the light cone. The massive Klein-Gordon equation has

$$G_4^F(\sigma;m^2) = iG_4^E(\sqrt{\sigma}) = \frac{im}{8\pi}\frac{1}{\sqrt{\sigma}}H_1^{(1)}(im\sqrt{\sigma}) \ . \tag{294}$$

If we expand this expression in the vicinity of the singularity at $\sigma = 0$ we find that there is a logarithmic singularity in addition to the pole. The logarithm is a multi-valued function; to handle this we place a branch cut along the negative real σ-axis that is then approached from above, in accordance with the analyticity properties of the Feynman propagator. To make a long story short the crucial step is

$$\ln(-\sigma) = \begin{cases} \ln|\sigma| \\ \\ \ln|\sigma| + i\pi \end{cases} \tag{295}$$

depending on whether the interval is timelike or spacelike. The final result is

$$\begin{aligned} G_F^4(\sigma) = -\frac{1}{4\pi}\delta(\sigma) - \frac{m^2}{8\pi^2}\theta(-\sigma)\left(\frac{1}{2} + \frac{m^2}{2^2 \cdot 4} + \ ...\right) - \\ -i\left(\frac{1}{4\pi^2}\frac{1}{\sigma} + \frac{m^2}{16\pi^2}\ln|\sigma| + \ ...\right) \ . \end{aligned} \tag{296}$$

This time the real part has support inside as well as on the light cone.

The situation in odd dimensions provides an interesting contrast. For the "massless" wave equation

$$G_3^F = -\frac{i}{4\pi}\frac{1}{\sqrt{\sigma}} \ . \tag{297}$$

Just like the massive propagator in four dimensions this Feynman propagator has a branch cut along the negative real axis, and the support of the commutator and retarded Green functions extends into the light cone.

From a practical point of view it makes a lot of difference whether the support of the commutator Green function extends into the light cone or not. If the support is confined to the light cone the result is that the effect of a sharply localized disturbance in the field will be seen as a distinct flash, or heard as a distinct sound if we are considering the characteristic cone of a sound wave. If the support extends into the cone effects will "ring on" for a considerable time, just like the waves on the two dimensional surface of a lake into which a stone has been thrown. Therefore talk is possible in a 3+1 dimensional spacetime, but not in a 2+1 dimensional one. The property that the support is confined to the characteristic cone is sometimes called "Huygens' principle in its strong form". It could also be taken as the definition of "massless", in which case there are no massless fields in odd dimensional spacetimes.

The conformally coupled scalar in $\mathbf{adS}_4$

When we go from Minkowski space to anti-de Sitter space three issues arise: Anti-de Sitter space has closed timelike curves, it is not globally hyperbolic (there is no Cauchy hypersurface), and it is not geodesically convex (there are pairs of points in that cannot be connected by geodesics). The first issue can be avoided by going to the universal covering space, but one can always ask whether it is necessary to do so. To deal with the second issue we have to take account of the possibility that influences may enter spacetime from spatial infinity, or else we may try to formulate boundary conditions that explicitly forbid this. The third issue means that the geodesic distance between two points may not exist, so that the Green function certainly cannot quite be a function of geodesic distance only. Actually they will be functions of

$$u = -X_1 \cdot X_2 \ . \tag{298}$$

However, the range of this variable will be over all the real numbers. There will then be three cases:

$$\begin{array}{ccl} u > 1 & \Rightarrow & u = \cosh d \ , \quad d \text{ spacelike distance} \\ |u| < 1 & \Rightarrow & u = \cos d \ , \quad d \text{ timelike distance} \\ u < 1 & \Rightarrow & \text{geodesic distance undefined} \ . \end{array} \tag{299}$$

An obvious question is: Should the Green function be zero when $u < 1$?

It is advisable to begin our anti-de Sitter deliberations by a discussion of the conformally coupled scalar in $\mathbf{adS}_4$ since this ought to behave like the massless scalar in Minkowski space—and if so it is a particularly simple case. The Feynman propagator obtained by analytical continuation from $\mathbf{H}^4$ (with Dirichlet conditions on the boundary) is

$$G_4^F = -\frac{i}{4\pi^2}\left(\frac{1}{(X_1 - X_2)^2} + \frac{1}{(X_1 + X_2)^2} \right) \ . \tag{300}$$

Remarkably there are two singular points inside anti-de Sitter space, given by

$$X_1^\alpha = \pm X_2^\alpha \ . \tag{301}$$

Compared to hyperbolic space, what has happened is that the "image charge" has moved into spacetime. Indeed this is what one can expect from the geometry of anti-de Sitter space. Timelike geodesics emerging from the first of the singular points will reconverge at the second. So will lightlike geodesics provided that they are reflected by $\mathcal{J}$, and then a second singularity in the Green function there is unavoidable. Because of the temporal periodicity of the Green function the closed timelike curves in anti-de Sitter space pose no problem; if we nevertheless decide to work in covering space we have to use some coordinate system that allows us to do so and we will find that there are an infinite number of periodically recurring singularities, not just two of them.

The analyticity properties of the Feynman propagator must be discussed with some care. The propagator should be analytic in the second and fourth quadrant of complex sausage time. Alternatively let the source point be at $U = 1$, $V = 0$, and continue analytically in the coordinates for the observation points according to

$$V = -iV_E \ . \tag{302}$$

We move the pole off the real axis according to

$$V \rightarrow V - \text{sgn}V i0 \ . \tag{303}$$

Then we find that

$$(X_1 - X_2)^2 = 2u - 2 + i0 \ . \tag{304}$$

The Green function must therefore be analytic in the upper u half plane. Since the singularities are poles the real part of the Feynman propagator is found to be

$$\bar{G}_4 = -\frac{1}{4\pi}\delta\left((X_1 - X_2)^2\right) - \frac{1}{4\pi}\delta\left((X_1 + X_2)^2\right) \ . \tag{305}$$

The support of this Green function—and hence of the commutator and retarded Green functions—is confined to the light cone, so Huygens' principle holds in its strong form. In particular the commutator Green function is

$$\tilde{G}_4 = \tilde{\epsilon}(t)\bar{G}_4 \ , \qquad \tilde{\epsilon}(t) = \text{sgn}(\sin t) \ , \tag{306}$$

where t is the sausage time. An analysis of the commutator Green function confirms our intuition of how the signal behaves: It emerges from the singular point $X_1^\alpha = X_2^\alpha$ (say), reaches $\mathcal{J}$, and then bounces back and reaches the second singular point along its backwards light cone. Since the singularities are of equal strength the entire signal is reflected—nothing comes in from across the boundary.

To clarify what goes on at $\mathcal{J}$ it is instructive to study the Green functions of the conformally coupled scalar on **adS** in the conformally compactified space—in this case on "one half" of the Einstein universe where $\mathcal{J}$ can be regarded as an ordinary timelike hypersurface. Choose sausage coordinates to describe anti-de Sitter space. The metric $\hat{g}_{ab}$ in the Einstein universe is related to the anti-de Sitter metric g_{ab} through

$$\hat{g}_{ab} = \Omega^2 g_{ab} \ , \qquad \Omega = \frac{1-\rho^2}{1+\rho^2} \ . \tag{307}$$

The Green function $\hat{G}$ in the Einstein universe is

$$\hat{G}(1,2) = \Omega^{-1}(1)G(1,2)\Omega^{-1}(2) \ . \tag{308}$$

We place the origin $t = \rho = 0$ at the source point, and then we find

$$(X_1 \mp X_2)^2 = -2 \mp \frac{1-\rho^2}{1+\rho^2} = -2 \mp \Omega^{-1}\cos t \ . \tag{309}$$

Let us concentrate on the Hadamard function that solves the homogeneous wave equation. As a result of a minor calculation we get

$$\hat{G}^{(1)} = \frac{1}{4\pi^2}\left(\frac{1}{\cos t + \Omega} - \frac{1}{\cos t - \Omega}\right) \ . \tag{310}$$

Interestingly this vanishes at $\mathcal{J}$, where $\Omega = 0$. In effect we have imposed Dirichlet conditions at the boundary, which explains why the signal came bouncing back from there.

In the hyperbolic case we were able to choose between various boundary conditions at the conformal boundary, and this is the case here as well. The difference is that changing the boundary conditions at $\mathcal{J}$ will change the strength of the singularity at the image point—and this makes a difference because the image point is situated inside **adS**. Specifically, consider

$$G_4^F = -\frac{i}{4\pi^2}\left(\frac{1}{(X_1 - X_2)^2} - \frac{k}{(X_1 + X_2)^2}\right) \ . \tag{311}$$

Clearly the cases $k = \pm 1$ are singled out by the requirement that the two singularities should have the same strength—if this is not the case we must conclude that the signal is not reflected in its entirety at $\mathcal{J}$, rather influences are entering or leaving spacetime there. In the conformally related Einstein universe we find that the Hadamard function becomes

$$\hat{G}^{(1)} = \frac{1}{4\pi^2}\left(\frac{1}{\cos t + \Omega} + \frac{k}{\cos t - \Omega}\right) \ . \tag{312}$$

There are two choices of boundary conditions that are totally reflective: Dirichlet conditions

$$\hat{G}^{(1)}_{\Omega=0} = 0 \quad \Rightarrow \quad k = -1 \tag{313}$$

and Neumann conditions

$$\partial_\Omega \hat{G}^{(1)}_{\Omega=0} = 0 \quad \Rightarrow \quad k = 1 \ . \tag{314}$$

This is not to say that other choices are inconsistent. Indeed if we take the point of view that we want to study radiation entering or leaving spacetime the totally reflective boundary conditions must be avoided. On the other hand if we really want to stay in **adS** with its closed timelike curves (rather than going to its covering space) then the reflective boundary conditions are necessary.

General anti-de Sitter Green functions in four dimensions

The conformally coupled scalar is the analogue of massless fields in Minkowski space. In four dimensions we were able to avoid the question of what the lack of geodesic convexity means for the Green functions, since the support of the commutator Green function was confined to the light cone and therefore does not extend into the region that is separated by a spacelike distance from the position of the image charge. For massive fields on the other hand this issue has to be faced squarely. It is clear that if the support extends into this

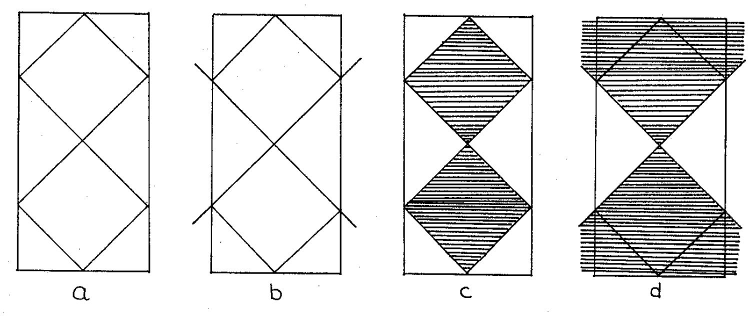

Figure 11: Some supports for the commutator Green function. a: A massless scalar with reflective boundary conditions. b: A generic massless scalar. c: A massive scalar for special values of μ^2. d: A generic massive scalar.

region then the conclusion must be that the radiation is leaking out through $\mathscr{J}$, and if we want to stay in anti-de Sitter space (with its closed timelike curves) this must be avoided.

It turns out that the precise value of μ^2 is crucial here. There are three interesting ranges for the variable u: When $u > 1$ the geodesic distance between the observation point and the "original" source point is spacelike, when $|u| < 1$ it is timelike, and when $u < -1$ the two points cannot be connected by any geodesic. The question is whether the Green function is necessarily zero also in this third region. The Feynman propagator for a massive scalar field is

$$G_4^F(\sigma;\mu^2) = \frac{i}{4\pi^2}\partial_u Q_\nu(u) \ . \tag{315}$$

The parameter ν was given as a function of μ^2 when we discussed its analytical continuation in $\mathbf{H}^4$, and u is as usual. The Legendre functions are singular at $u = \pm 1$ and there is a branch cut along the real axis between the singular points. A detailed analysis shows that the support of the real part extends into the region where

$$-1 \leq u \leq 1 \ , \tag{316}$$

that is to say into those regions of **adS** that can be connected to the source point by means of timelike geodesics, so that the formula

$$\cos d = u \tag{317}$$

makes sense. This is as expected for a massive field. Generically the support also extends into the region where $u < -1$, which means that radiation is leaking out of spacetime. However, it can be shown that for special ("quantized") values of μ^2 this does not happen. It follows that these special values of the mass must be adopted in anti-de Sitter space proper. This is in agreement with what one would expect from group theory—since the mass has to do with the eigenvalue of the Killing vector J_{UV} (that generates time translations) and since this generates a compact subgroup of $SO(3,2)$ we do expect a quantization of mass to occur. On the other hand the relevant group for the universal covering space is the non-compact group of time translations, and the range of allowed μ^2 should be continuous. Unfortunately the details are an exercise in special functions.

Green functions in $\mathbf{adS}_3$

It is more instructive to discuss Green functions in $\mathbf{adS}_3$ where we do not have to delve into the Bateman manuscript—we know that the Green function on $\mathbf{H}^3$ can be expressed with elementary functions. By analytic continuation of the (Dirichlet) Green function on $\mathbf{H}^3$, we get the Feynman propagator

$$G_3^F(\sigma;\mu^2) = -\frac{i}{4\pi}\frac{e^{-\sigma\sqrt{\mu^2+1}}}{\sinh\sigma} \ . \tag{318}$$

This is for spacelike geodesic distances σ. Using our variable u we get for source and observation points in general position that

$$G_3(u;\mu^2) = -\frac{i}{4\pi}\frac{1}{\sqrt{u^2-1}}\frac{1}{(u+\sqrt{u^2-1})^{\sqrt{\mu^2+1}}} \ . \tag{319}$$

Now the function

$$\sqrt{u^2-1} = \sqrt{u-1}\sqrt{u+1} \tag{320}$$

has branch points at

$$u = \pm 1 \ . \tag{321}$$

There is a branch cut in between. Since we require analyticity in the upper half plane it is real and positive to the right of the cut, and real and negative to the left. Along the cut it is imaginary. If we set

$$\sqrt{\mu^2 + 1} \in \mathbf{Z} \tag{322}$$

there are no further complications. The Feynman propagator will develop a real part only in the region $|u| < 1$, and the support properties of the commutator Green function will be the same as that advertized for special values of μ^2 in $\mathbf{adS}_4$. On the other hand the conformally coupled scalar has

$$\sqrt{\mu^2 + 1} = \frac{1}{2} , \tag{323}$$

and in this case the support of the commutator Green function will extend into the region $u < -1$.

These matters can be further clarified by a study of the representation theory of the anti-de Sitter groups, but our story ends here.

LITERATURE

For the conformal properties of projective quadrics, and a good deal else besides, I recommend

- R. Penrose: Relativistic symmetry groups, in A.O. Barut: Group Theory in Non-Linear Problems, Reidel, Dordrecht 1974.

Chapters two and three are based on a series of papers by S. Åminneborg et al. that can be found in Classical and Quantum Gravity, or alternatively on gr-qc. In chapter four I followed

- A. Ashtekar and A. Magnon, Asymptotically anti-de Sitter Spacetimes, Class. Quant. Grav. **1** (1984) L39.

In particular they claim that it "is straightforward to show" that eq. (203) implies eq. (204). I failed to do it. For more and deeper results on this topic see

- H. Friedrich, Einstein's equation and geometric asymptotics, gr-qc/9804009, to appear in the GR15 proceedings.

For the Hamiltonian viewpoint (which has led to some interesting things that I originally meant to talk about) see

- J.D. Brown and M. Henneaux, Central charges in the canonical realization of asymptotic symmetries: An example from three dimensional gravity, Commun. Math. Phys. **104** (1986) 207.

A standard reference for Green functions in anti-de Sitter space is

- S.J. Avis, C.J. Isham and D. Storey, Quantum field theory in anti-de Sitter space-time, Phys. Rev. **D18** (1978) 3565.

For the group theory chapter I never wrote, I recommend

- A.O. Barut and C. Fronsdal, On non-compact groups II. Representations of the 2+1 Lorentz group, Proc. Roy. Soc. (1965) 532.