Introduction to Cosmology

ASTR 434

Andrew Cooper

apcooper@gapp.nthu.edu.tw

Lecture 2 / The Expanding Universe / September 24th 2024

This lecture: the expanding Universe

- The Hubble Flow and the Hubble Constant
- The Cosmological Principle
- Special and general relativity (review/information only, *not examined*)
- The Friedman-Lemaître-Robertson-Walker Metric
- Redshift as a measure of expansion
- Curvature (review/information only, *not examined*)

Recap

M33 NASA

Angular size and brightness depend on distance.

Objects of known brightness (or known size) can be used to infer distances.

For example, the luminosity of certain types of stars can be inferred from other information, such as their colour or pulsation period.

When this ‘standard candle’ method was used (in the 1920s) to infer distances to galaxies, two important things were discovered:

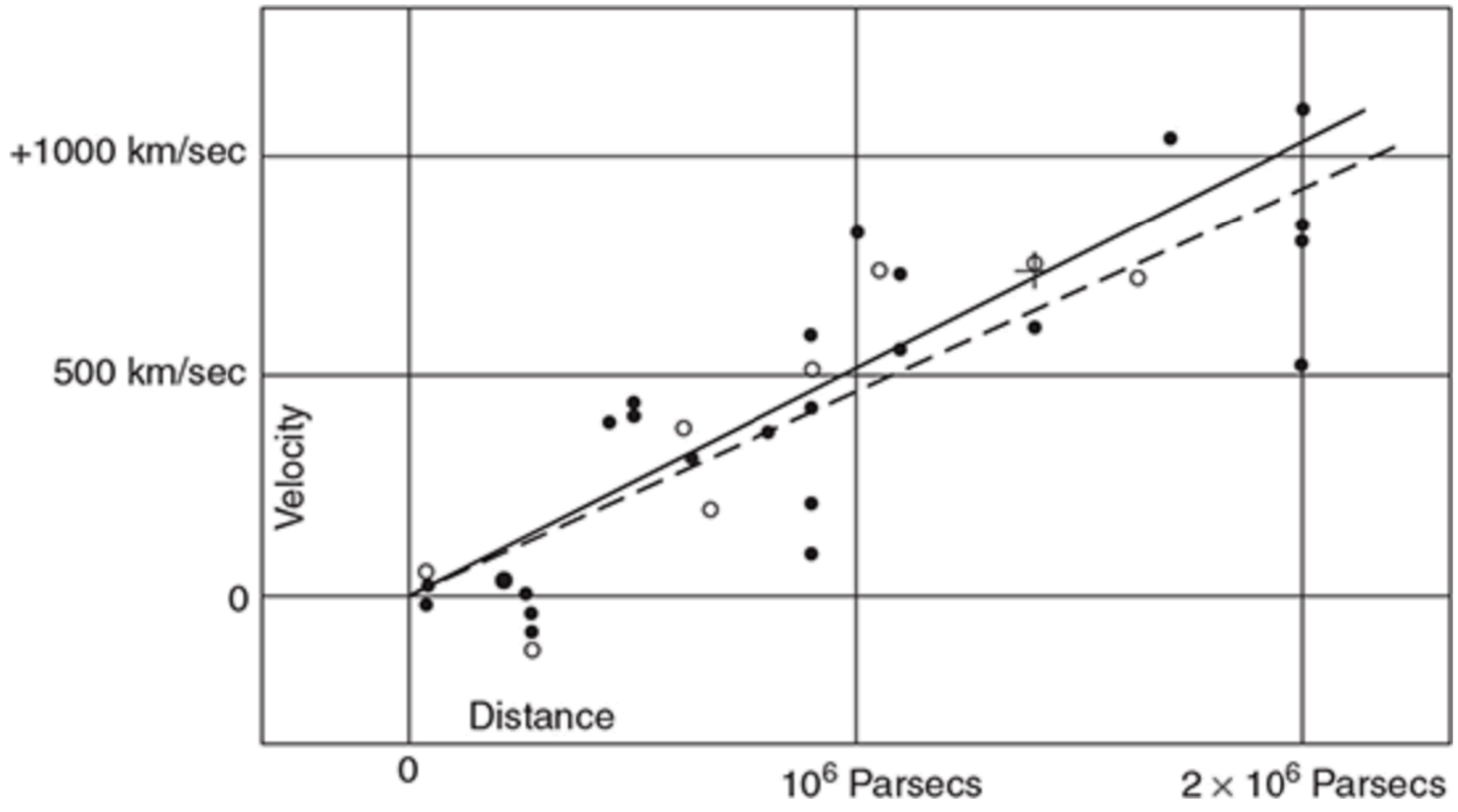
Galaxies are really very far away!

There is a (rather surprising) somewhat-linear relationship between the distance to galaxies and the ***redshifts*** *of their spectra (equivalently, their* ***apparent velocities*** *relative to the Sun).*

The Hubble Flow

The Hubble Flow

Adapted by Ryden from Hubble 1929 (PNAS 15, 168)



$v = cz$

"Hubble's Law"

$$\boxed{v = H_0 d}$$

$$\left(\text{or } z = \frac{H_0}{c} d\right)$$

$$d_L = \left(\frac{L}{4\pi F}\right)^{1/2}$$

Standard candle luminosity distances

Hubble Constant

H_0: the constant of proportionality in Hubble's law; units are $\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

Empirically, this is the speed with with galaxies **close to us today** appear to be **receding** in a **systematic expansion:** the **recession velocity**.

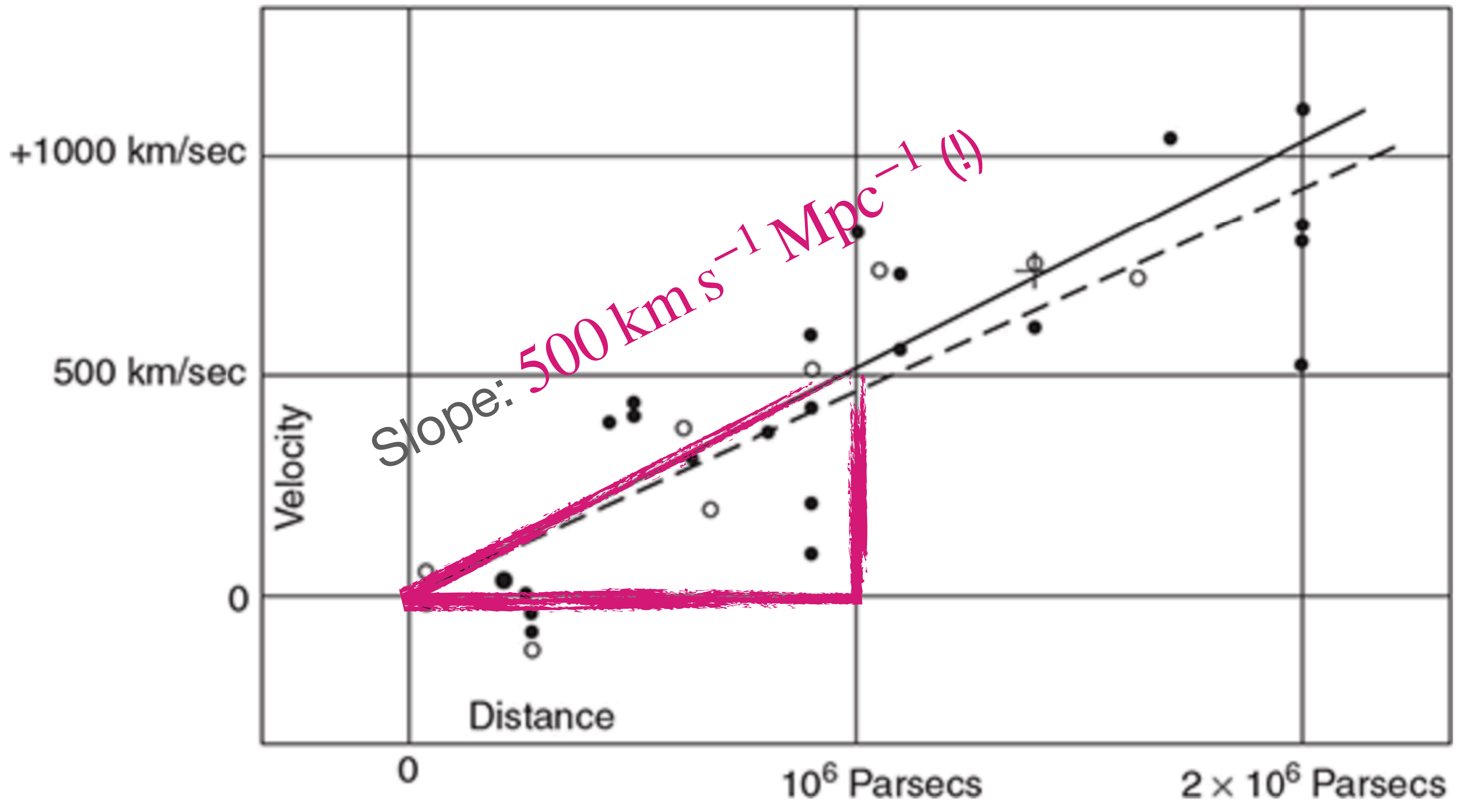
The Hubble Flow ($z \propto d$) is a **measurement,** i.e. an **observational fact** — *explaining* the Hubble Flow in terms of an expanding Universe is a **theoretical interpretation**, not something that is obvious from the measurement on its own.

The Hubble Constant is one of the fundamental **cosmological parameters**. It is central to our physical understanding of the history of the Universe.

In principle, H_0 can be **measured directly**: the more points (galaxies) we have on the Hubble diagram, the more accurate our measurement should be. However, things are not so easy.

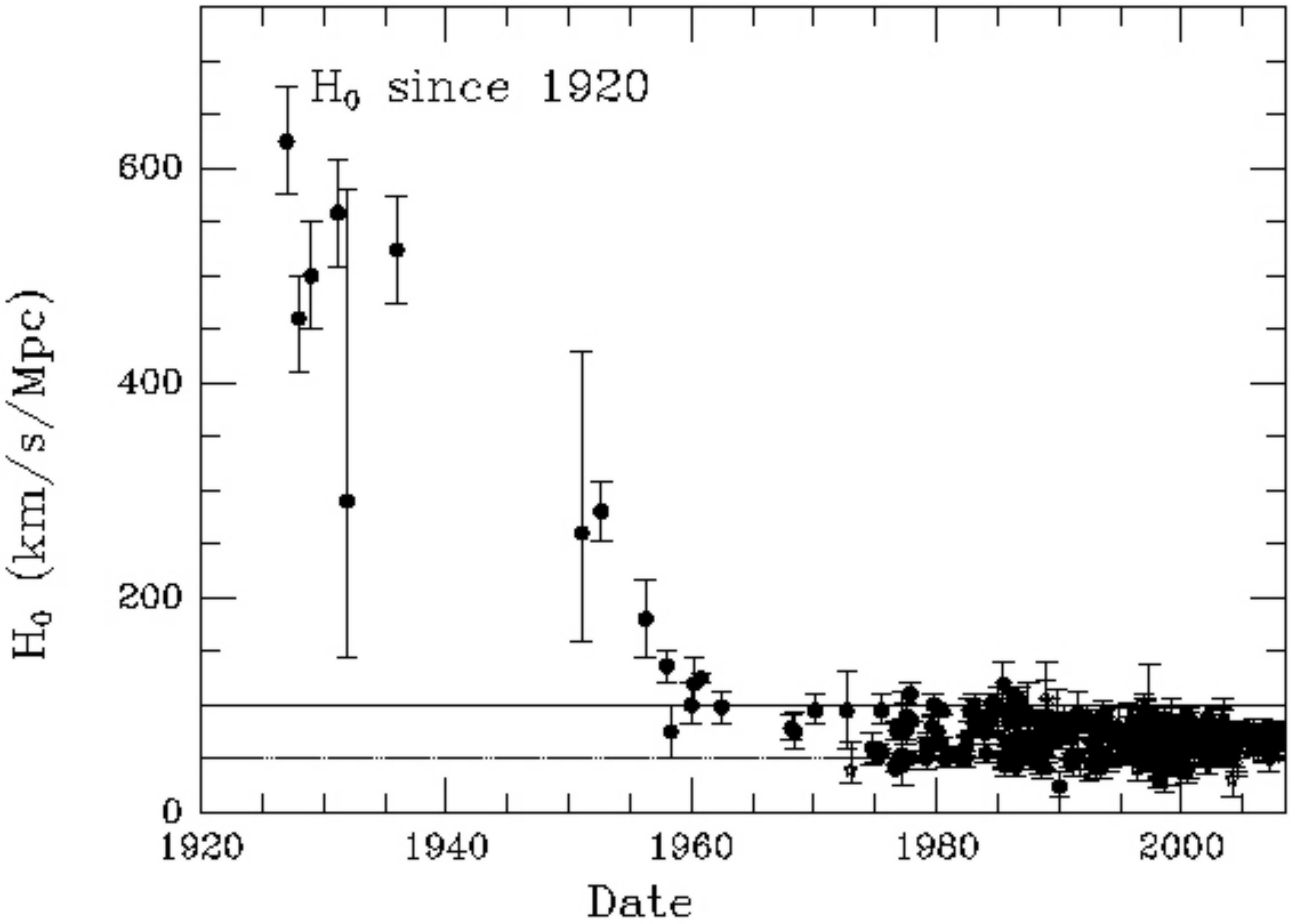
Hubble Constant

Adapted by Ryden from Hubble 1929 (PNAS 15, 168)



Hubble Constant

https://lweb.cfa.harvard.edu/~dfabricant/huchra/hubble/



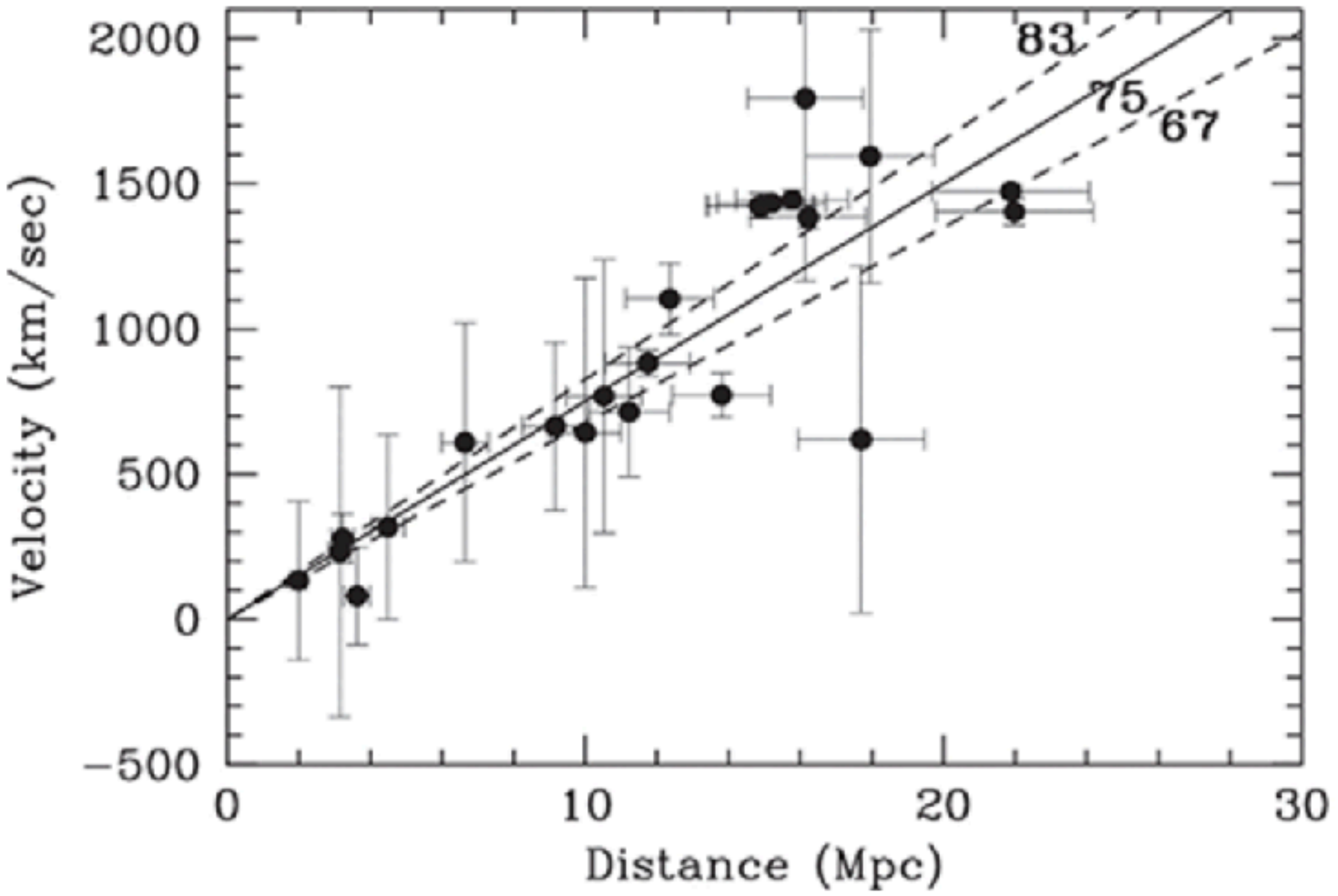
For a long time, H_0 was poorly constrained.

$$50 \lesssim H_0 \lesssim 100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}.$$

Rather than choose a particular value, cosmologists preferred to write $H_0 = h\,100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, with "little h" a dimensionless constant, $0.5 \lesssim h \lesssim 1$.

Hubble Constant

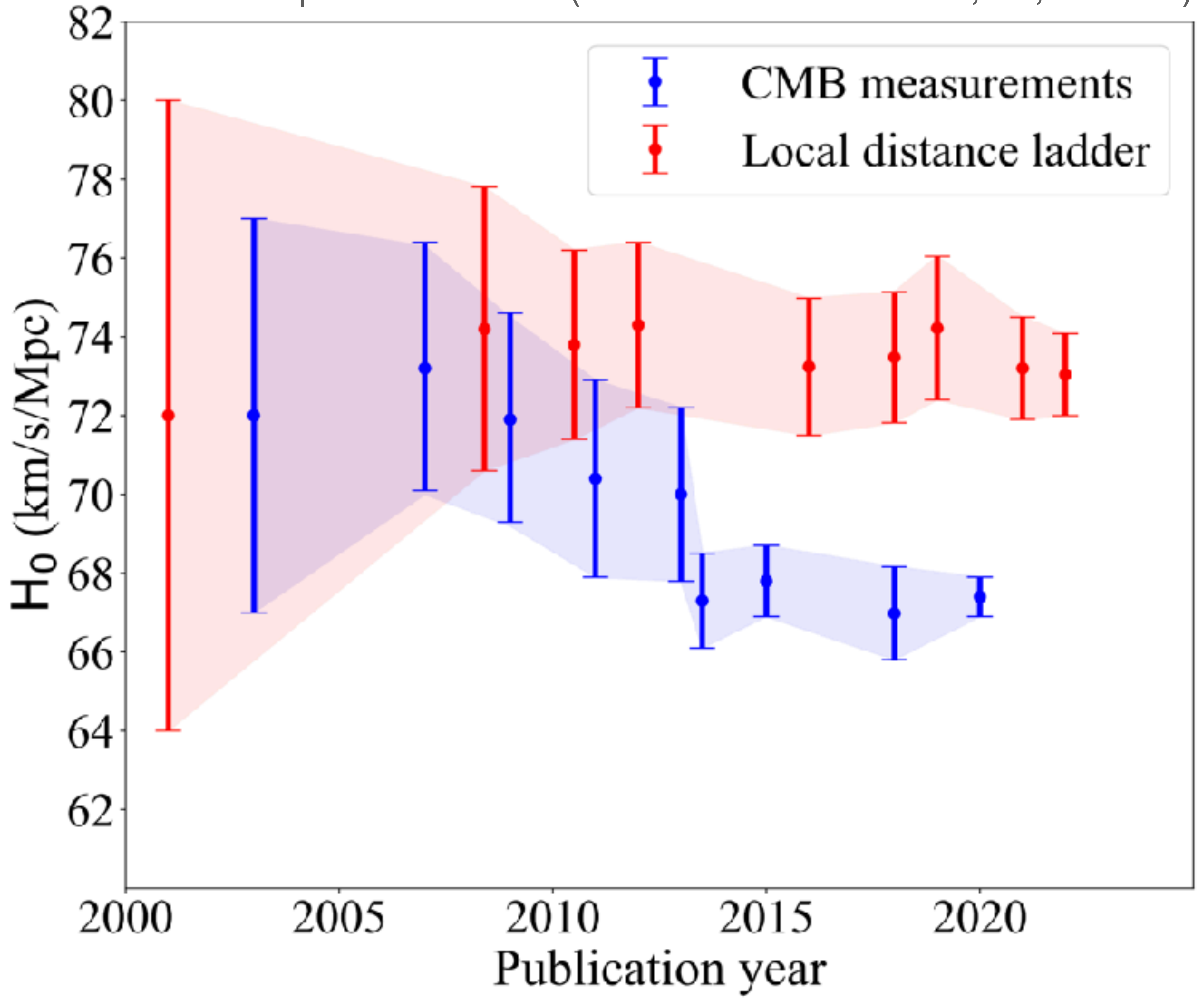
Adapted by Ryden from Freedman et al. 2001 (ApJ 553, 47)



Using Hubble's method, based on distances to nearby galaxies, $H_0 \simeq 73 \pm 1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$

Hubble Constant

Perivolaropoulos & Skara (New Astron. Rev. 2022, 95, 101659)



Using Hubble's method, based on distances to nearby galaxies, $H_0 \simeq 73 \pm 1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$

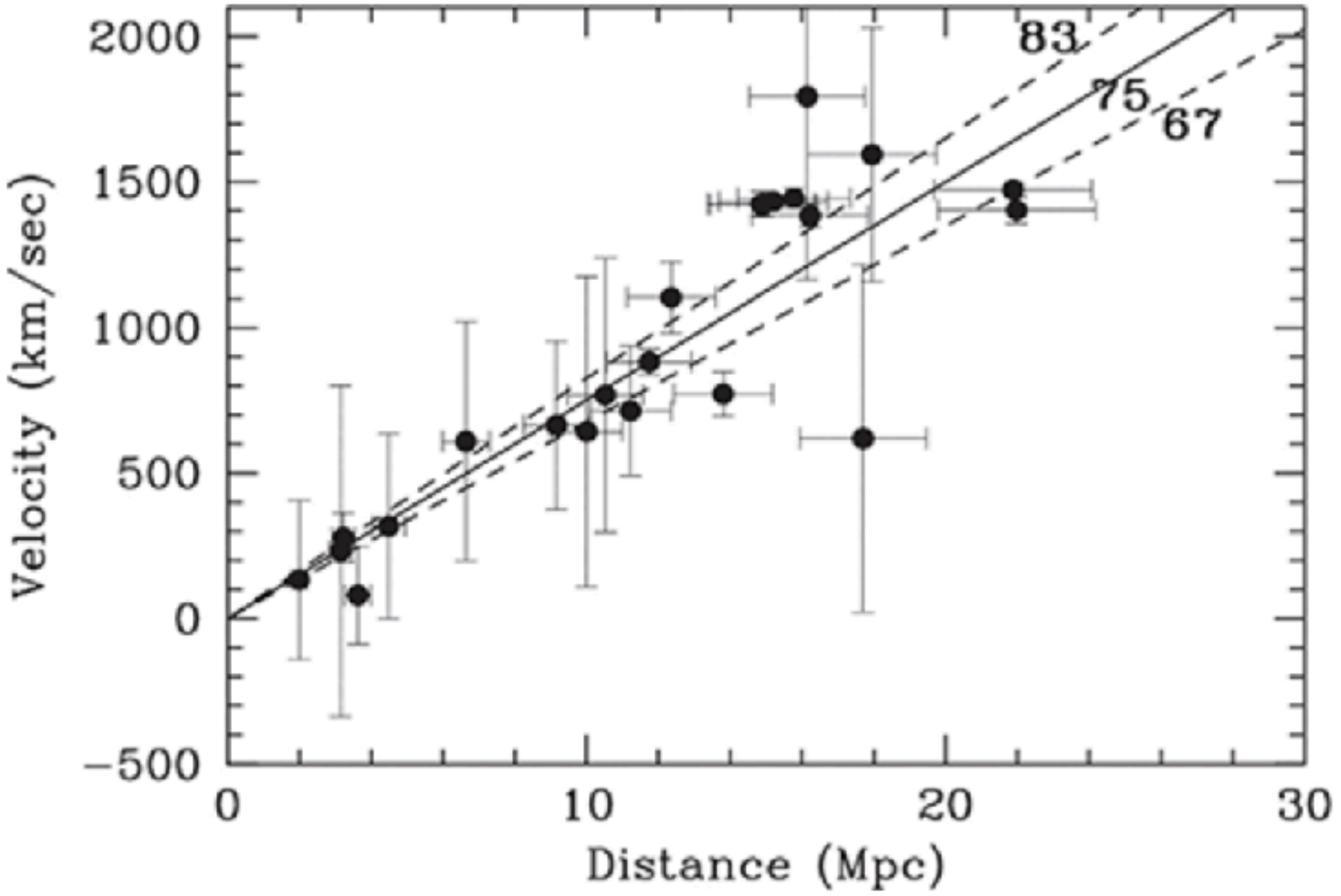
There is still disagreement between Hubble's method ("local") and a totally different way of measuring H_0 using the cosmic microwave background ($H_0 \simeq 68 \pm 2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.).

As far as we can tell, both measurements are correct, so cosmologists think this disagreement is very exciting. This is called the H_0 **tension**.

Hubble Constant

$H_0 = 68 \pm 2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

Adapted by Ryden from Freedman et al. 2001 (ApJ 553, 47)



<u>Scatter</u> around the Hubble Flow is due to **peculiar motion** (i.e. local gravitational accelerations). This does not go away with better measurements.

Unlike the Hubble flow, the peculiar motions don't vary in any systematic way with distance. As the distance increases, the "Hubble velocity" becomes much larger than the typical peculiar velocity.

Uncertainty in the local measurement of H_0 is due to uncertainty in distance measurements.

Hubble Time

Notice that the Hubble constant has dimension $[T]^{-1}$. The reciprocal of H_0 is therefore a time: the **Hubble Time.**

$$t_H \equiv 1/H_0 \sim \frac{1}{68\,\mathrm{km\,s^{-1}\,Mpc^{-1}}} \simeq 14.7\,\mathrm{Gyr}$$

Useful fact: $1\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \simeq 0.001\,\mathrm{Gyr^{-1}}$

The Hubble Time is a **characteristic timescale** for the expansion. It is an **estimate** of the time since all the points expanding with the Hubble Flow were "in the same place" — i.e. the time since the "Big Bang".

<u>But it is just an estimate</u>, because it assumes that separations have *always increases at the same rate,* H_0. As we will, see, this is not true.

Our best estimate of the age of the Universe is $\sim 1\,\mathrm{Gyr}$ less than t_H (the fact that it is not very different is something of a cosmic coincidence). Despite this, t_H is still a useful number in situations we will meet later.

Hubble Length

The distance at which the recession speed is equal to the speed of light defines a length scale, the Hubble Length:

$v = c = H_0 d_H$

$$\implies d_H(t_0) = \frac{c}{H_0} \sim 3000\,h^{-1}\,\mathrm{Mpc}$$

(the ~ 3000 is from $c = 299792.458\,\mathrm{km\,s^{-1}} \sim 3 \times 10^5\,\mathrm{km\,s^{-1}}$)

This is also the distance light can travel in a Hubble Time.

There is nothing really fundamental about the Hubble Length, it is just a characteristic length scale. The combination c/H_0 shows up in a lot of the equations we will work with later.

We see plenty of galaxies with apparent recession velocities larger than c. On the scale approaching the Hubble Length, pretending that $v = cz$ (or even the special-relativistic version of that) becomes an increasingly bad approximation.

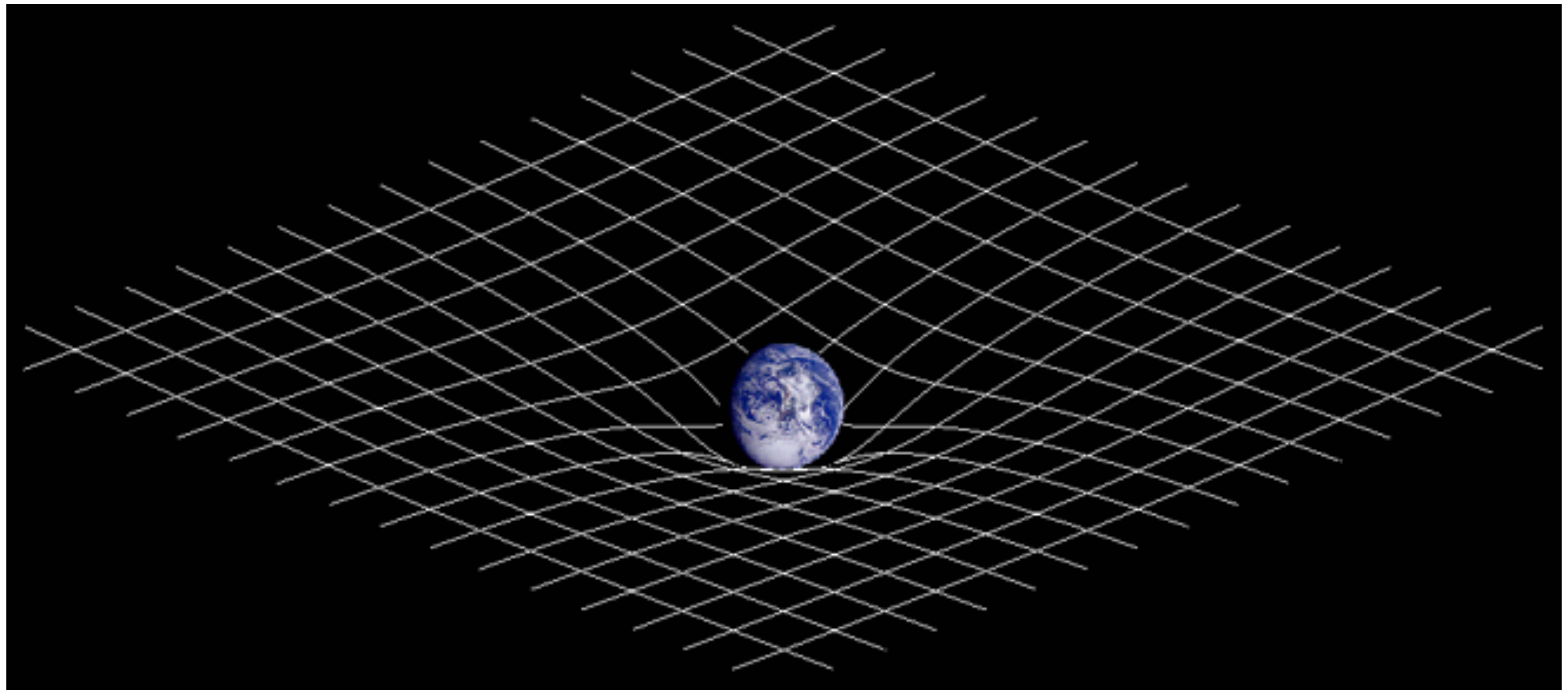
Expanding Spacetime

Dynamic spacetime

The **expansion of space** is a prediction of General Relativity (GR).

In GR, "spacetime" is *dynamic*: it changes in response to the local density of mass / energy ($E = mc^2$).

This idea is often illustrated with a "ball on a rubber sheet" picture, like this:



Source unknown

Dynamic spacetime

We will see that the Hubble Flow can be explained by applying the principles of GR to “the whole universe”, i.e. to an infinite, uniform distribution of mass.

GR introduces two key “dynamical” ideas about space that are fundamental to cosmology: **expansion** and **curvature**.

To apply GR to any problem, we need to specify **how the mass-energy is distributed** and choose an appropriate **metric**.

These are all complicated ideas, so we will start with some basic concepts and work back slowly to the application of GR in cosmology.

What do we mean by “space”?

We are familiar with the idea of **static Euclidean** space. In such a space, we can represent positions of objects with respect to a **fixed grid of reference points** (the coordinate grid).

In a 1-dimensional Euclidean space we need one coordinate, x, to specify a position. In 2-dimension space we need a pair of coordinates, (x, y), and so on.

These might be *Cartesian* coordinates (x, y) or *polar* coordinates (r, θ) (in 3-d we have cartesian coordinates or spherical polar coordinates, or some mix of the two).

What do we mean by “space”?

We can find the shortest distance between two points (for example, in 3 dimensions) using Pythagoras’ theorem (here in Cartesian coordinates):

$$s_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2.$$

Usually, we think of “objects” (galaxies) as moving, i.e. changing their position with respect to the “coordinate grid” over time.

In this picture, the distances between the coordinate grid points do not change: two objects *at rest* with respect to each other will always be the same “coordinate” distance apart.

What do we mean by “space”?

It doesn’t have to be that way. The coordinate grid (“space”) itself can change with time.

In the case of the Hubble flow, we can ignore (for simplicity, for now) the small peculiar motions, and hence pretend the galaxies are **locally at rest** (i.e. that they don’t acquire any significant velocity due to gravitational acceleration over a Hubble time).

This is the same as saying each galaxy is sitting at rest on some grid point $\mathbf{x} \equiv (x_1, x_2, x_3) \equiv (r, \theta, \phi)$ of a coordinate system.

In cosmology, we call the grid points “galaxies” — this is just a lazy convention. If we like, we could call them “fixed comoving observers”.

Expanding space

If we just want to *describe* the Hubble flow, without explaining anything, we can introduce a coordinate system that changes with time (in the same way everywhere). In such a system, *all* the galaxies move away from each other over time *without any local accelerations*.

Like this:

$$\mathbf{x}_{\text{phys}} = a(t)\,\mathbf{x}_{\text{comoving}}$$

$\mathbf{x}_{\text{phys}}$ is the "physical" position of a galaxy (grid point). In an expanding coordinate system, this increases over time.

$\mathbf{x}_{\text{comoving}}$ is the "comoving" position of a galaxy. This does not change with time. We need to choose a reference time to define the *comoving separation* between galaxies.

$a(t)$ is the ***scale factor*** (or ***expansion factor***) of the coordinate system, which is the part that depends on time. By definition, the time t_0 when $a(t_0) = 1$ fixes the comoving separation between galaxies. We are free to choose this time.

The function $a(t)$ could be anything! This is where we will need to put in the physics of the expansion.

Expanding space

https://lickrishvillage.com/light-and-sweet-bread/

Expanding space



https://lickrishvillage.com/light-and-sweet-bread/

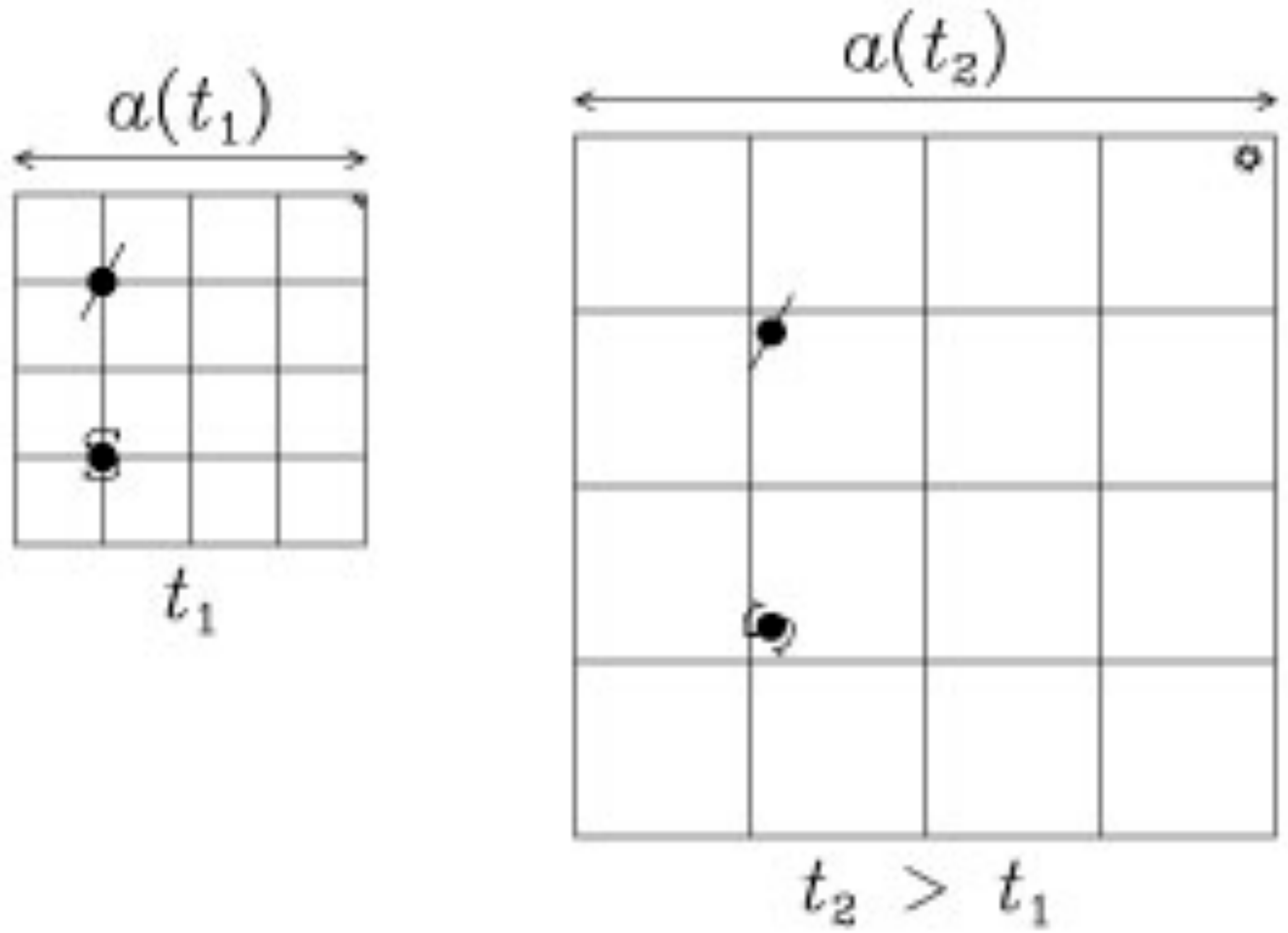
Expanding space

$$\mathbf{x}_{\rm phys} = a(t)\, \mathbf{x}_{\rm comoving}$$

Often physical and comoving coordinates are distinguished with different letters rather than subscripts (although there is not much agreement on the choice of letters).

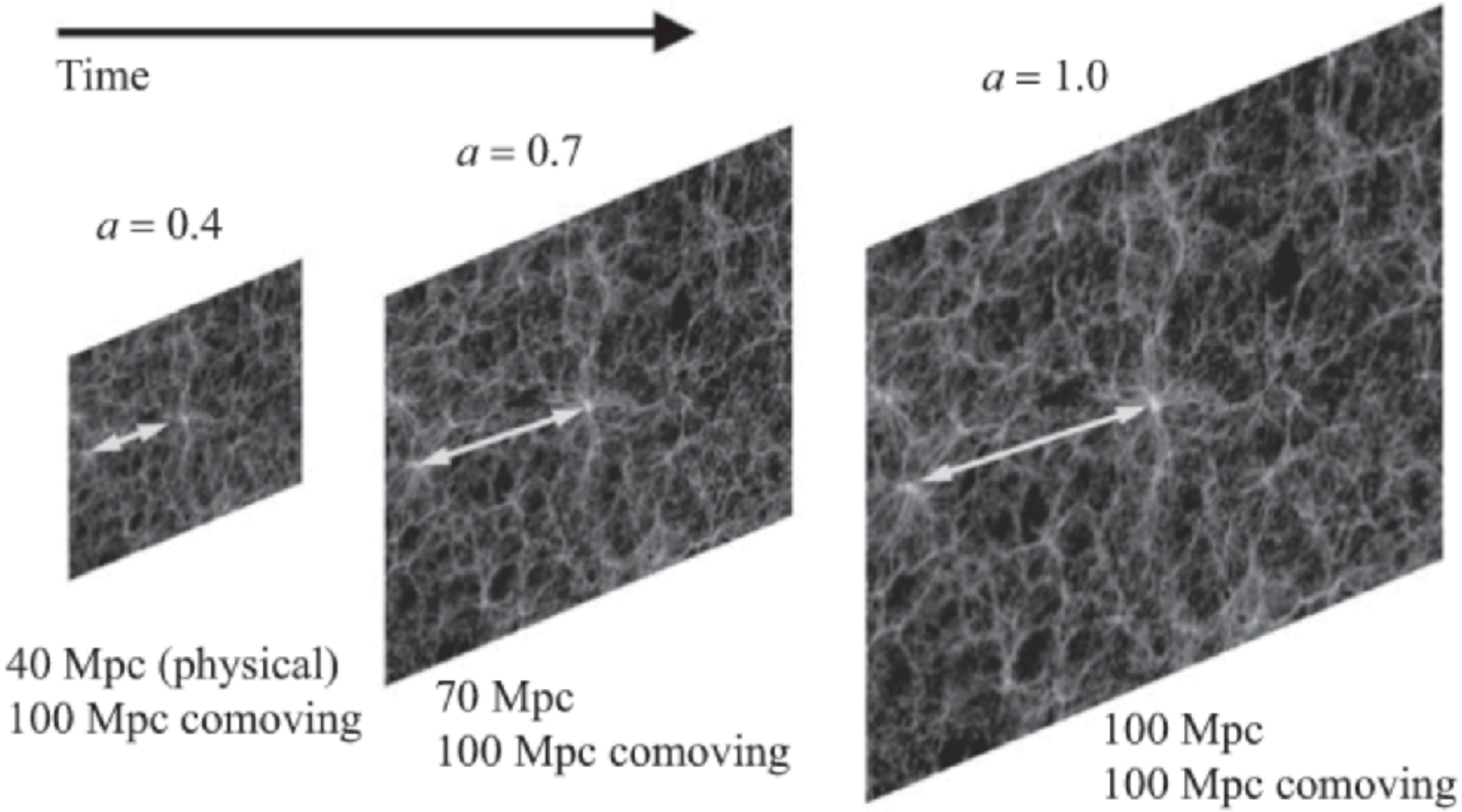
I will use **x** for physical coordinates and **r** for comoving coordinates:

$$\mathbf{x} = a(t)\, \mathbf{r}$$



E. Bertschinger

https://ned.ipac.caltech.edu/level5/March02/Bertschinger/Bert_contents.html



From Huterer

Expansion factor

For example, two galaxies have comoving positions $\mathbf{r}_1 = (3,0,0)$ and $\mathbf{r}_2 = (0,4,0)$ at $t = t_0$. At this time, their separation is $s = 5$ units.

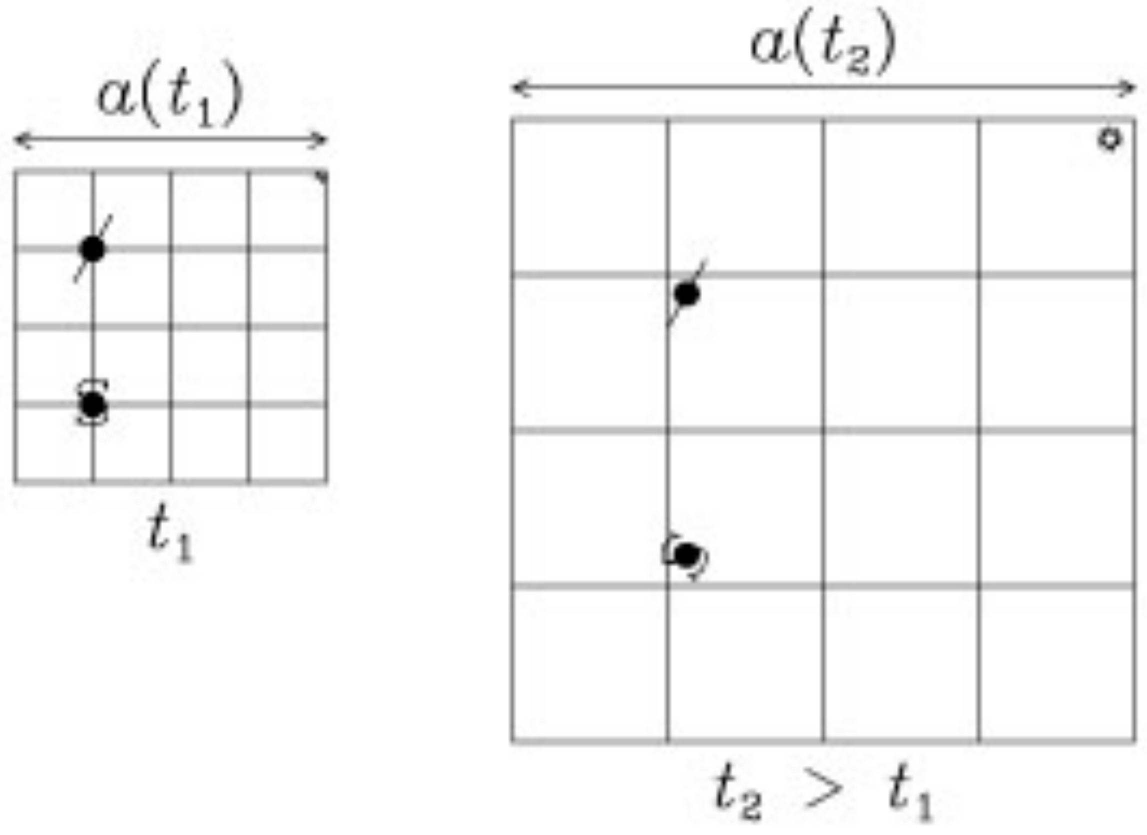
Let's imagine $a(t) = 10t$, a linear expansion history with a constant expansion rate ($\dot{a} = 10$).

Taking arbitrary times $t = 1$, $t = 2$, $t = 4$ we have $a(1) = 10$, $a(2) = 20$, $a(4) = 40$; i.e. when the time coordinate doubles, the scale (distance between the galaxies) also doubles.

More generally, the ratio of expansion factors at different times is :

$$\frac{a(t)}{a(t_0)} = \frac{10t}{10t_0} = \frac{t}{t_0}$$

so $a(t) = \dfrac{t}{t_0}$ if we define $a(t_0) = 1$.



Expansion factor

The comoving separation of the points is always the separation at $t = t_0$, $s = 5$ units.

At $t = 2t_0$ we have $a(t) = \dfrac{2t_0}{t_0} = 2$, so $\mathbf{x} = a(2t_0)\,\mathbf{r} = 2\,\mathbf{r}$:

$\Longrightarrow \mathbf{x}_1(t = 2t_0) = (6{,}0{,}0)$, $\mathbf{x}_2(t = 2t_0) = (0{,}8{,}0)$ and hence $s_{21}(t = 2t_0) = \sqrt{6^2 + 8^2} = 10$ units.

The distance between the points at $t = 2t_0$ is twice the distance between them at $t = t_0$, because the coordinate scale has expanded.

All distances scale up as $a(t)$ — after all, coordinates are just distances from the origin.

If we fix “ourselves” to be the origin and look at another galaxy, we can write $s_{\text{phys}} = a(t)\,s_{\text{comoving}}$ where s_{phys} is the physical distance between us and the other galaxy.

Expansion factor

There is nothing too exotic about introducing an expanding coordinate system.

We can do this in any situation where we want to separate out relative motion and a background expansion. Cosmology is just one application of this technique.

Only ratios of expansion factors matter, so we are free to define $a = 1$ at some reference time for mathematical convenience.

In cosmology we almost always use t**he present day (**$t = t_0$**)** as the reference time.

The expansion history

One of the goals of cosmology is to figure out $a(t)$ for the real Universe. We call this the <u>expansion history</u>.

We will see that we can use GR to predict $a(t)$, **<u>if</u>** we can first specify some fundamental properties of the Universe.

We don't know the values of those properties for the real Universe.

We proceed by *measuring* the actual $a(t)$ as well as we can (i.e. **observing**) and *inferring* those values, by finding the theoretical $a(t)$ that matches what we observe.

The question now becomes: what numbers do we need to know, and how do we predict $a(t)$ from those numbers?

Expansion rate

What can we observe? Well, we can start with the derivative of $a(t)$, called the **expansion rate**:

$$\frac{d}{dt}s_{\text{phys}} \equiv v(t) = \frac{da}{dt}s_{\text{comov}} = \dot{a}\, s_{\text{comov}}$$

(remember $\frac{d}{dt}s_{\text{comov}} = 0$ by definition, and we have $\dot{a} \equiv \frac{da}{dt}$ to save writing)

So we have the rate of change of physical distance between galaxies (comoving observers) with time, due to the expansion: $v(t) = \dot{a}s_{\text{comov}}$

In terms of their physical separation of the galaxies at time t, $v(t) = \frac{\dot{a}}{a}s_{\text{phys}}$.

This should look familiar: it is the velocity of comoving objects seen by a comoving observer.

The Hubble Parameter, $H(t)$

$$v(t) = \frac{\dot{a}}{a} s_{\text{phys}}$$

has the same meaning as Hubble's Law:

$$v = H\, d_{\text{phys}}$$

We can therefore identify $H(t) \equiv \dfrac{\dot{a}}{a}$. This is a general mathematical result, nothing to do with the specific physics of the expansion.

We have been careful to write the **Hubble Parameter** $H(t)$ as a function of time. At $t = t_0$, $H(t_0) = \dfrac{\dot{a}(t_0)}{a(t_0)} = H_0$, the **Hubble Constant** (remember $a(t_0) = 1$).

So, in an *expanding coordinate system*, the Hubble Constant $H_0 \equiv \dot{a}(t_0)$ **is the present-day rate of expansion**, which we can measure directly.

Summary: At the present day, **all** galaxies (comoving observers) see other galaxies moving away from them, at a rate proportional to their distance. H_0 is the constant of proportionality. In the past or the future, we might observe a different $H(t)$.

Expansion factor and Hubble Parameter

$H(t) = \dot{a}/a$ is a very important tool for manipulating cosmological models and connecting them to observations.

You should remember it, and be able to explain what it means.

Remember, there is no deep physics in the equation $H(t) = \dot{a}/a$: it is just describing an expanding coordinate system.

The actual function we use for $H(t)$ will follow directly from a function for $a(t)$.

Note: in general, $0 < a \leq 1$, because $a = 1$ at the present day.

Light signals

We learn about the Universe using light signals, which take a finite time to propagate, and do so at a fixed speed c.

This introduces a major headache into our perspective on the Universe, but also some benefits — we can see into the past, and use that information to constrain the expansion history.

"Events" are coordinates in *spacetime*, a 4-dimensional coordinate system consisting one one time coordinate and three spatial coordinates; for example (t, x, y, z).

The "events" (galaxies) we observe "now" are those on our **past lightcone**; the time and distance at which we observe the events is determined by how long it takes light to propagate between them and us.

Light signals in Special Relativity

Special relativity deals with how different ***inertial*** (non-accelerating) observers perceive the same events, depending on the relative velocity of their local frames of reference. Special relativity cannot deal with expanding or curved spacetime.

The machinery of special relativity is built around the idea that there are certain things that different inertial observers must agree on, one being the speed of light, and another being the "laws of physics".

These are called "invariants".

A consequence of invariance of the laws of physics is that the distance *in spacetime* between two events must be invariant (distance does not depend on the choice of coordinates).

If these things are invariant, *something else* must vary between observers with different frames of reference.

$\Longrightarrow$ Inertial observers in relative motion will **disagree** on *distances in space* (i.e. lengths) and *distances in time*, when those things are treated separately. Rules get shorter or longer, clocks tick slower or faster.

Minkowski spacetime metric

We can define a distance in (static, Euclidean) spacetime using the *Minkowski metric*:

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

In this metric, there is one "time" dimension and three "space" dimensions.

For a signal propagating along the radius vector between the source and the observer (length $r^2 = x^2 + y^2 + z^2$):

$$ds^2 = -c^2dt^2 + dr^2$$

Light is the limiting case of a signal for which $ds = 0$. You can see this makes some sense in terms of the metric, because, in that case, we have:

$$\Delta t = \frac{\Delta r}{c}$$

i.e. the separation in time between two events is equal to the light travel time between the space positions of those two events. A path for which $ds = 0$ is called a **null geodesic**.

General Relativity

The Minkowski metric of special relativity applies to inertial motion in a **static** universe; no accelerations, so no gravity. Such a Universe is said to have **Euclidean** geometry.

General relativity introduces a second-order coupling between coordinates in the metric: **curvature of spacetime**.

This curvature scales with the density of mass-energy. “Gravity” is the consequence of motion along trajectories in a curved spacetime.

Various interesting consequences, including:

- Black holes
- Gravitational lensing
- Gravitational waves

These are usually treated in the case of **strong** effects due to locally high concentrations of mass-energy. However, the idea that “mass curves spacetime” applies also to the Universe as a whole, since there is *some* mass-energy *everywhere*.

The cosmological principle

What do we "know" about the mass distribution of the Universe? What can we safely assume in order to make progress?

Rough version: "The universe is infinite and the same everywhere. There are no special places in the Universe".

More precise version, the **cosmological principle**: "***The Universe is homogeneous and isotropic***".

Homogeneous: the same everywhere.

Isotropic: appearing the same in every direction.

A Universe that is isotropic around *every point* is necessarily homogeneous.

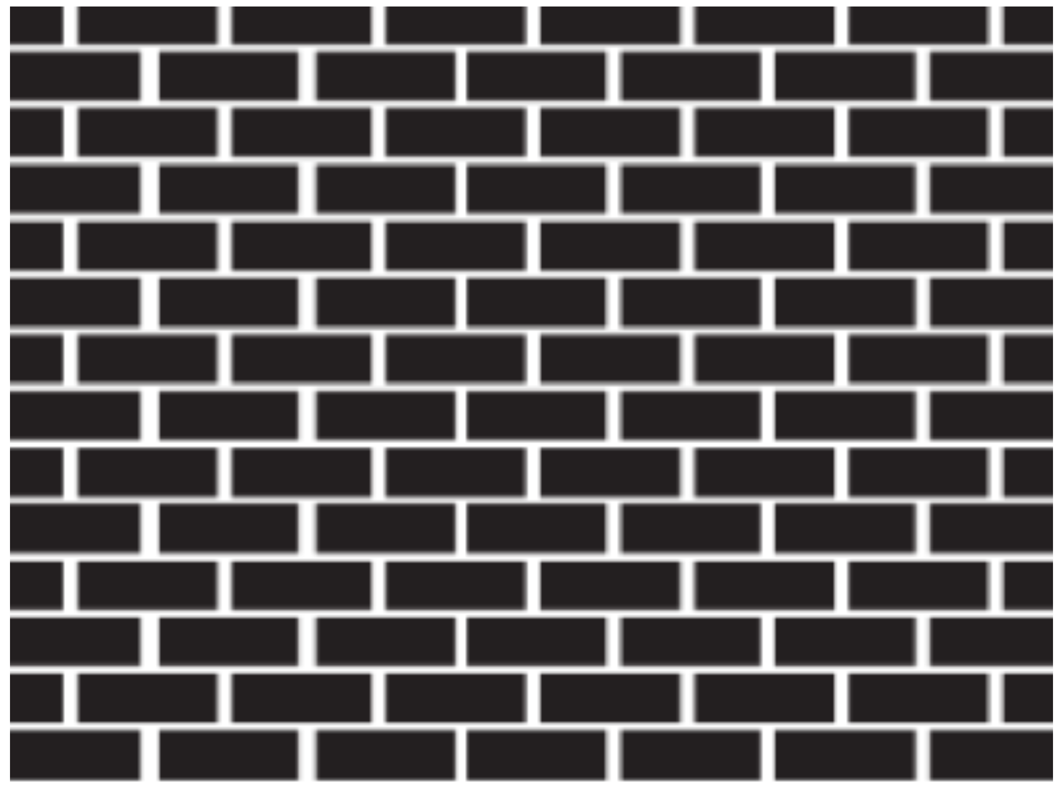
This does not *have* to be true; as we saw last week, it *appears* to be *approximately* true and that approximation *appears* to get better on larger and larger scales.

On small scales it is not true, so all the following **only applies on large scales.**

Examples of homogeneity and isotropy

Homogeneous (on large scales), but not isotropic.

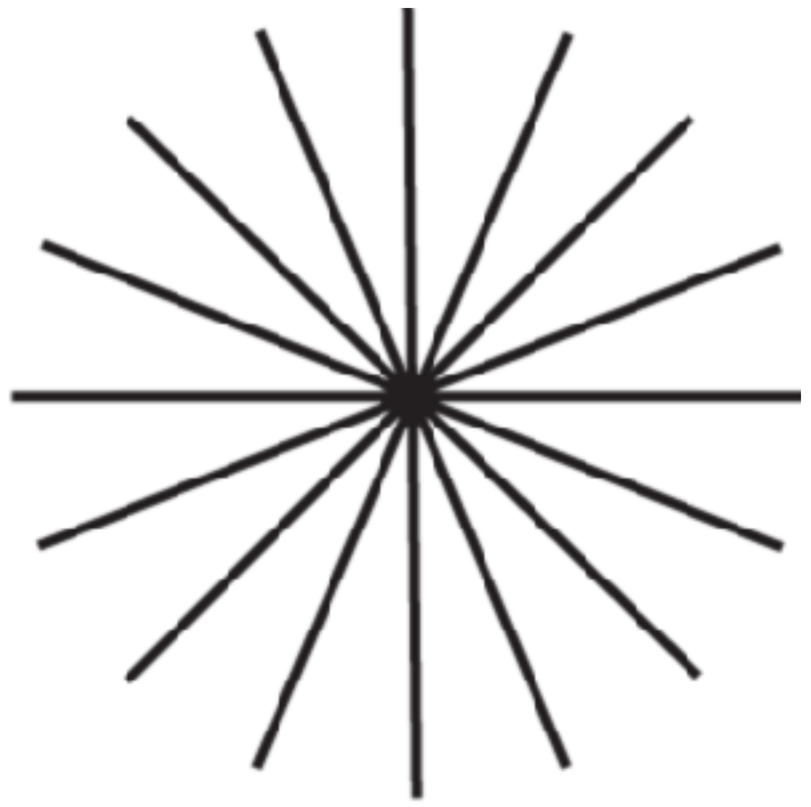
Ryden



Huterer

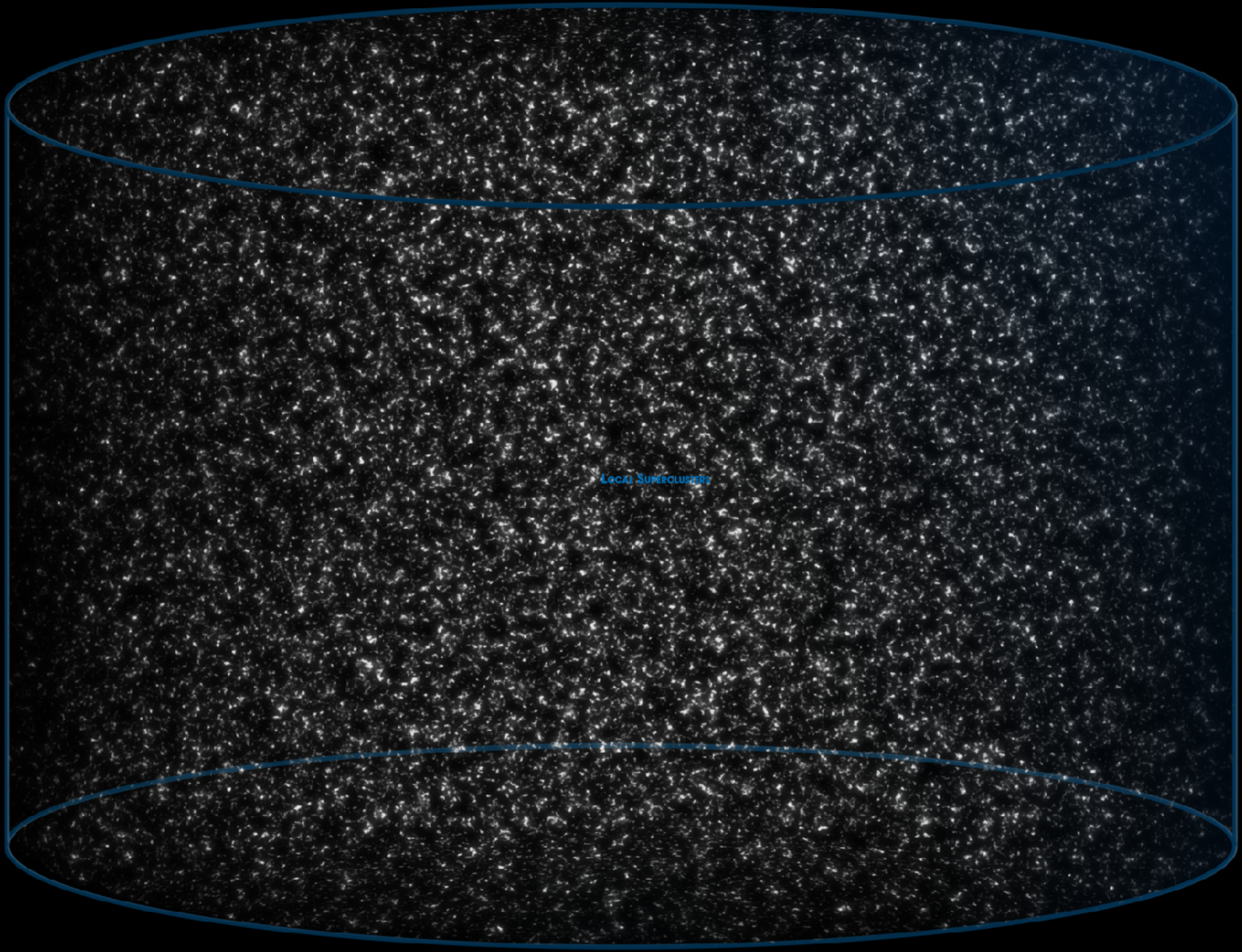
Isotropic (about a point), but not homogeneous.

Ryden



Huterer

OBSERVABLE UNIVERSE



Andrew Z. Colvin

Friedmann-Lemaître-Robertson-Walker metric

The FLRW metric is the most general metric for a homogeneous, isotropic mass distribution in 3-dimensional space.

Alexander Friedmann developed GR solutions for a homogeneous, isotropic mass distribution in the 1920s.

Lemaître recognized the cosmological significance of these solutions, in particular the expansion.

The modern, most general form of metric compatible with the cosmological principle was developed independently by Robertson and Walker.

FLRW Metric: the plan

The FLRW metric is an important general tool that provides a starting point for solving lots of cosmology problems.

The principles behind the FLRW metric are obviously important, but at this stage, the most important things are knowing **what it is** and **what the point of it is** — i.e. its **practical function** in the simple cases that we will apply it to.

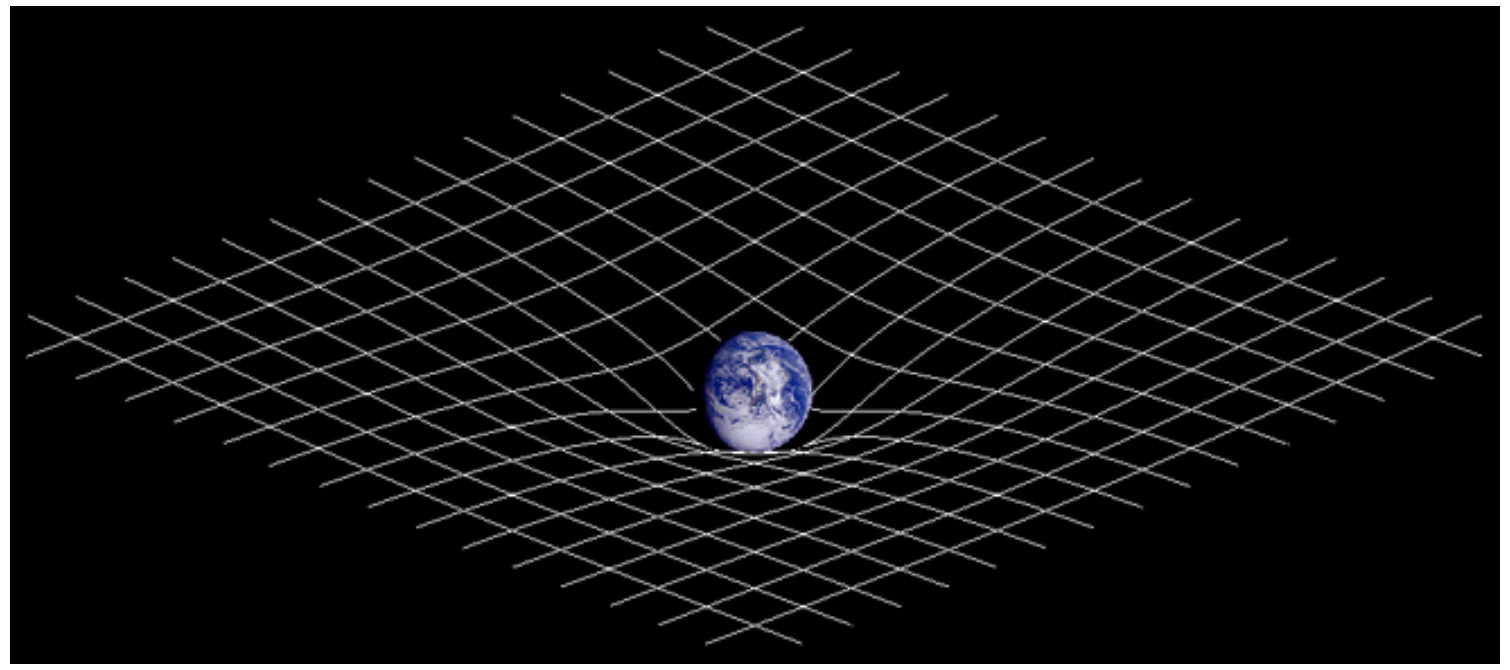
Therefore, we will learn about the different "bits" of FLRW in a slightly different order to the textbooks.

We will jump to the "main point" first, and look the special case of the metric that we will use most often.

Afterwards, we will go back to look at more "complicated" aspects that we will not use very often in the course.

(c) Victorinox

FLRW Metric: concepts



FLRW generalises the Minkowski metric to include the possibility of **expansion** and non-zero spatial curvature.

Unlike the textbooks, we will leave the details of what **“spatial curvature”** until ***after*** we look at the basic FLRW metric. For now, here is the short version:

We already said GR involves some sort of “bending” of the space coordinate system: this is **curvature**.

In the rubber-sheet picture, the mass is all in one place, so different parts of the sheet have different curvature. In an infinite mass distribution, the curvature is **everywhere**; in a homogeneous distribution the curvature must be **uniform**, i.e. **the same everywhere**. For example, the surface of a sphere is a 2-d surface with uniform curvature.

Fine, but we are dealing with the curvature of 3-dimensional space, not a 2-d rubber sheet! At this point the rubber sheet picture becomes rather useless (even misleading). It is impossible to draw a useful picture of “curved space” in 3-d!

FLRW Metric: three cases

Curvature changes the physical length of paths between nearby points in space.

If the spatial curvature must be uniform, the metric only has to deal with three cases:

1. No curvature (“flat”);

2. Positive curvature;

3. Negative curvature.

We will put (2) and (3) in a box for now.

As far as we know, the real Universe is case (1): flat. Since we are only doing “introduction to cosmology” we will mostly assume (1), and treat (2) and (3) as of only “academic interest”.

To appreciate the deeper idea of cosmology (and understand how we know the real Universe has no curvature) we will still take a quick look at (2) and (3) later.

FLRW Metric: simplest version

The <u>specific</u> FLRW metric for a <u>flat Universe</u>

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2]$$

Usually we don't need to distinguish the two transverse (angular) coordinates, so it's simpler to write the differential element of solid angle as $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$:

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + r^2d\Omega^2]$$

In this equation, a point in spacetime is specified by a coordinate with 4 elements: (t, r, θ, ϕ); r is the **<u>comoving</u> radial coordinate** that we introduced above. The angles are measured **on the celestial sphere (e.g. the right ascension and declination of galaxies**).

The spacetime displacement ds is **physical**. The time coordinate is also physical — it is the "cosmic clock".

You can see this is just the Minkowski metric in spherical polar coordinates, except for the time dependence introduced through the scale factor $0 < a(t) < 1$.

Proper distance

To get a feel for how the FLRW metric is used, we will compute the **proper distance** between two points in spacetime (i.e. between two galaxies).

The proper distance is the distance you would measure with a (very long) ruler.

In a static spacetime, galaxy A and galaxy B are always the same ruler-distance apart, so it doesn't matter *when* the measurement of distance is made.

In an expanding spacetime, the comoving separation is constant, but the physical separation changes. We can imagine *instantly* measuring the separation with our ruler at some particular time t.

Of course, this measurement is effectively *impossible* on cosmological scales in real life.

Anyway, let's calculate it.

Proper distance

We start from the metric:

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + r^2d\Omega^2]$$

We consider a separation along the radial direction ($d\Omega = 0$) at fixed time ($dt = 0$), so the metric simplifies to:

$$ds = adr$$

Then we integrate along the path between the two galaxies **in comoving coordinates**:

$$s_{\text{proper}} = \int ds = a \int_0^r dr' = ar$$

That was easy! We already met this distance in our discussion of Hubble's law (where we called it s_{phys}).

Redshift

Redshift: $z = \dfrac{\lambda_0 - \lambda_e}{\lambda_e} \implies 1 + z = \dfrac{\lambda_0}{\lambda_e}$ for light that is emitted with wavelength λ_e and received by the observer at λ_0.

We say **cosmological redshift is <u>not</u> a Doppler shift** because cosmological redshift is caused by the effect of the expansion on the propagation of light signals through spacetime. This has nothing to do with the rest frame motion of the source or the observer.

Speaking somewhat loosely, we often say that photon wavelengths are "stretched" by the expansion.

Redshift and scale factor

The redshift of an object due to the expansion of the Universe is directly related to the scale factor by the following equation:

$$a(t) = \frac{1}{1+z}$$

Equivalently,

$$1 + z = \frac{1}{a}$$

These equations are important and you need to remember them. Redshift is something we observe, $a(t)$ is something we predict.

On the next slides we will show how this follows from the FLRW metric.

Redshift and scale factor

For light, $ds^2 = 0$ (this is such a useful thing!)

$$cdt = a(t)dr$$

So if we integrate from the time when the light signal is emitted, t_e, to when we receive it, t_0, we have:

$$c\int_{t_e}^{t_0} \frac{1}{a(t)}dt = \int_0^r dr' = r$$

The time between peaks of a light wave is the frequency $\nu = c/\lambda$, so if we detect one peak at t_0, the next one will arrive at $t_0 + \lambda_0/c$. In the emission frame, the first peak was emitted at t_e and the second peak was emitted at $t_e + \lambda_e/c$. So for the second peak we can write:

$$c\int_{t_e+\lambda_e/c}^{t_0+\lambda_0/c} \frac{1}{a(t)}dt = r$$

Redshift and scale factor

These two peaks have travelled the same comoving distance, so:

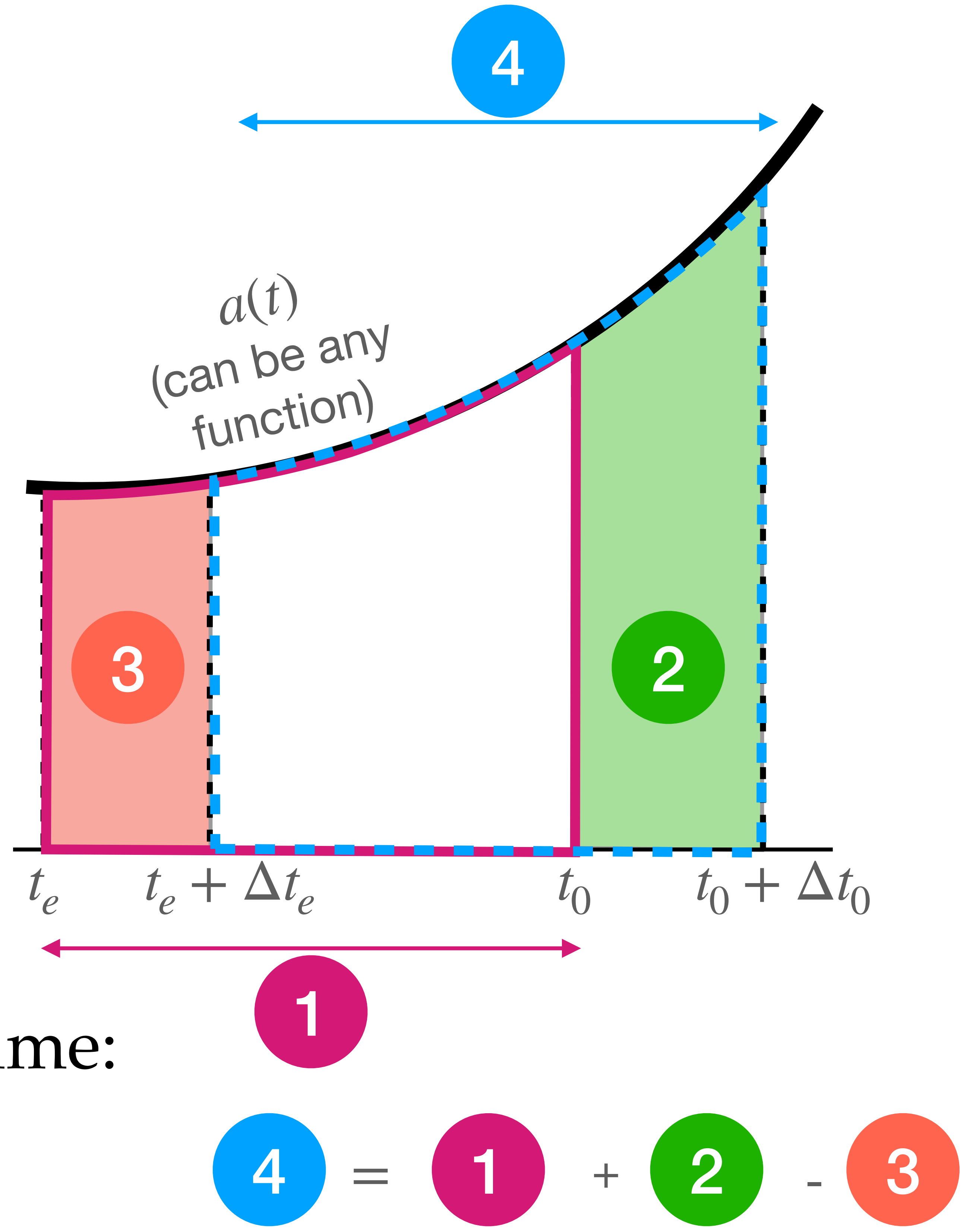
$$\int_{t_e}^{t_0} \frac{1}{a(t)} dt - \int_{t_e+\lambda_e/c}^{t_0+\lambda_0/c} \frac{1}{a(t)} dt = 0$$

Since the integrand is the same we can write the integrals as one operation:

$$\left[\int_{t_e}^{t_0} - \overset{4}{\boxed{\int_{t_e+\lambda_e/c}^{t_0+\lambda_0/c}}}\right] \frac{1}{a(t)} dt = 0$$

The 'trick' — we can rearrange the second integral into **three different ranges** of time:

$$\left[\underbrace{\int_{t_e}^{t_0}}_{1} - \overset{4}{\boxed{\left(\underbrace{\int_{t_e}^{t_0}}_{1} + \underbrace{\int_{t_0}^{t_0+\lambda_0/c}}_{2} - \underbrace{\int_{t_e}^{t_e+\lambda_e/c}}_{3}\right)}}\right] \frac{1}{a(t)} dt = 0$$



Redshift and scale factor

Rearranging,

$$\int_{t_e}^{t_e+\lambda_e/c} \frac{1}{a(t)} dt = \int_{t_0}^{t_0+\lambda_0/c} \frac{1}{a(t)} dt$$

If the time between peaks is short enough that $a(t)$ is approximately constant, i.e. $\delta t = \lambda/c \ll t$, then we have:

$$\frac{\delta t_e}{a(t_e)} = \frac{\delta t_0}{a(t_0)} \Longrightarrow \frac{\lambda_0}{\lambda_e} = \frac{a(t_0)}{a(t_e)} = \frac{1}{a(t_e)}$$

Since $1 + z = \lambda_0/\lambda_e$, it follows that:

$$1 + z = \frac{1}{a(t)}$$

for a light signal emitted at expansion factor $a(t)$.

Redshift and scale factor

We can use $a(t)$, $z(t)$ or just t as equivalent labels of "time" in cosmology. Only $z(t)$ is observable.

The redshift tells us about how much the Universe has expanded since the signal was emitted; nothing more or less!

Whether we use $a(t)$, $z(t)$ or t depends on the problem we're trying to solve. In casual conversation, cosmologists (and astronomers) prefer to use redshift.

$a(t)$ and $z(t)$ are interchangeable, but converting either of them to "clock time" requires a specific function for $a(t)$.

Note that $a = 1 \implies z = 0$. We very often refer to the present-day Universe as "redshift zero".

Geometry and curvature

The following slides review the concept of curvature and the associated notation that appears in the general form of the FLRW metric, and also the Friedman equations we will talk about next week.

You do not need to reproduce any of this in the homework or the exam.

Geometry and curvature

In a space of two dimensions we can find the shortest distance between two cartesian coordinates (x, y), $(x + dx, y + dy)$, using Pythagoras' theorem:

$$ds^2 = dx^2 + dy^2$$

If we use polar coordinates (r, θ) instead of (x, y), this **metric** becomes

$$ds^2 = dr^2 + r^2 d\theta^2$$ (can be shown by substituting x and y written as functions of r and θ)

A displacement in the radial direction is not equivalent to a change in the polar angle θ; one is a length and the other is an angle!

When combining the displacements to get a total distance, the angular displacement has to be multiplied by the radius at which that displacement occurs. This makes sense: the path around the circumference of a circle is longer when the radius is larger.

The total displacement ds does not depend on the choice of coordinates.

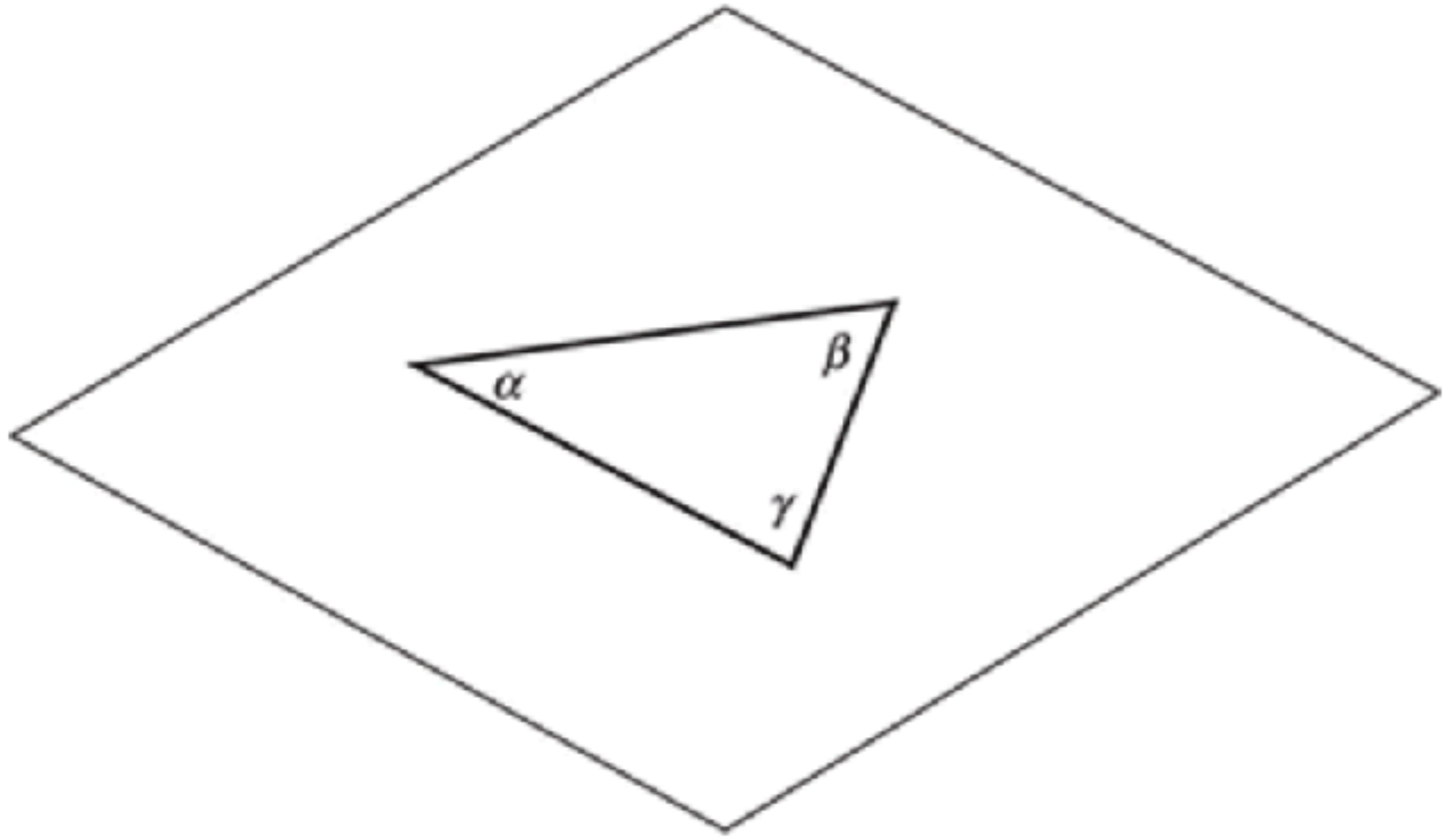
Flat space

Triangles and circles are shapes that live on a 2-dimensional surface.

Some familiar results from high-school geometry:

The angles in a triangle add up to $180°$ *(*π *radians).*

The circumference of a circle is given by $C = 2\pi r$*, where* r *is the radius.*



However, these results are **not true in general** for all two-dimensional surfaces. They only apply to triangles and circles drawn on a **flat surface** — not all 2-dimensional surfaces are flat!

Surface of sphere

The surface of a 3-dimensional sphere is 2-dimensional.

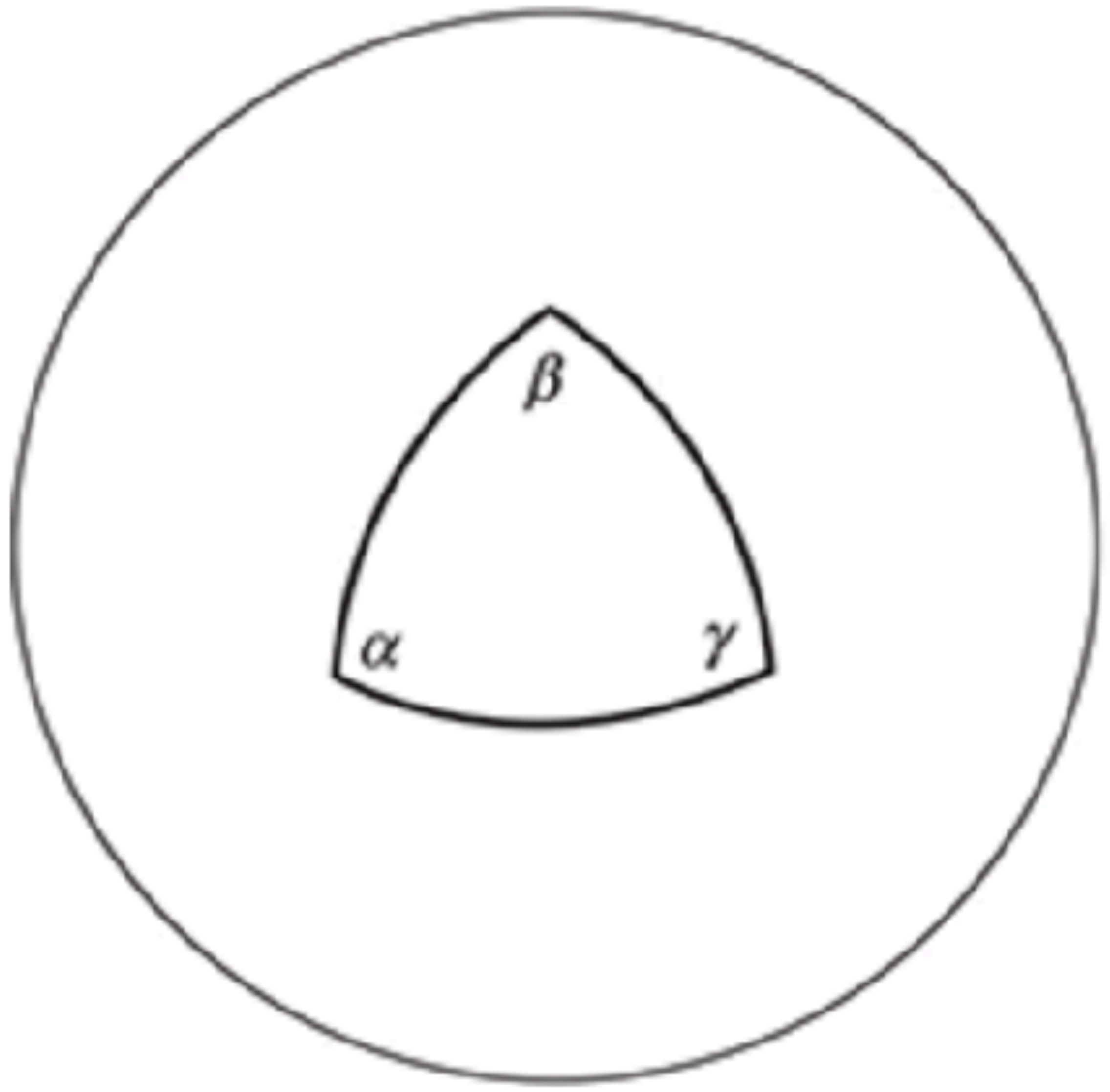
On the surface of a sphere of radius R, the angles in a triangle add up to:

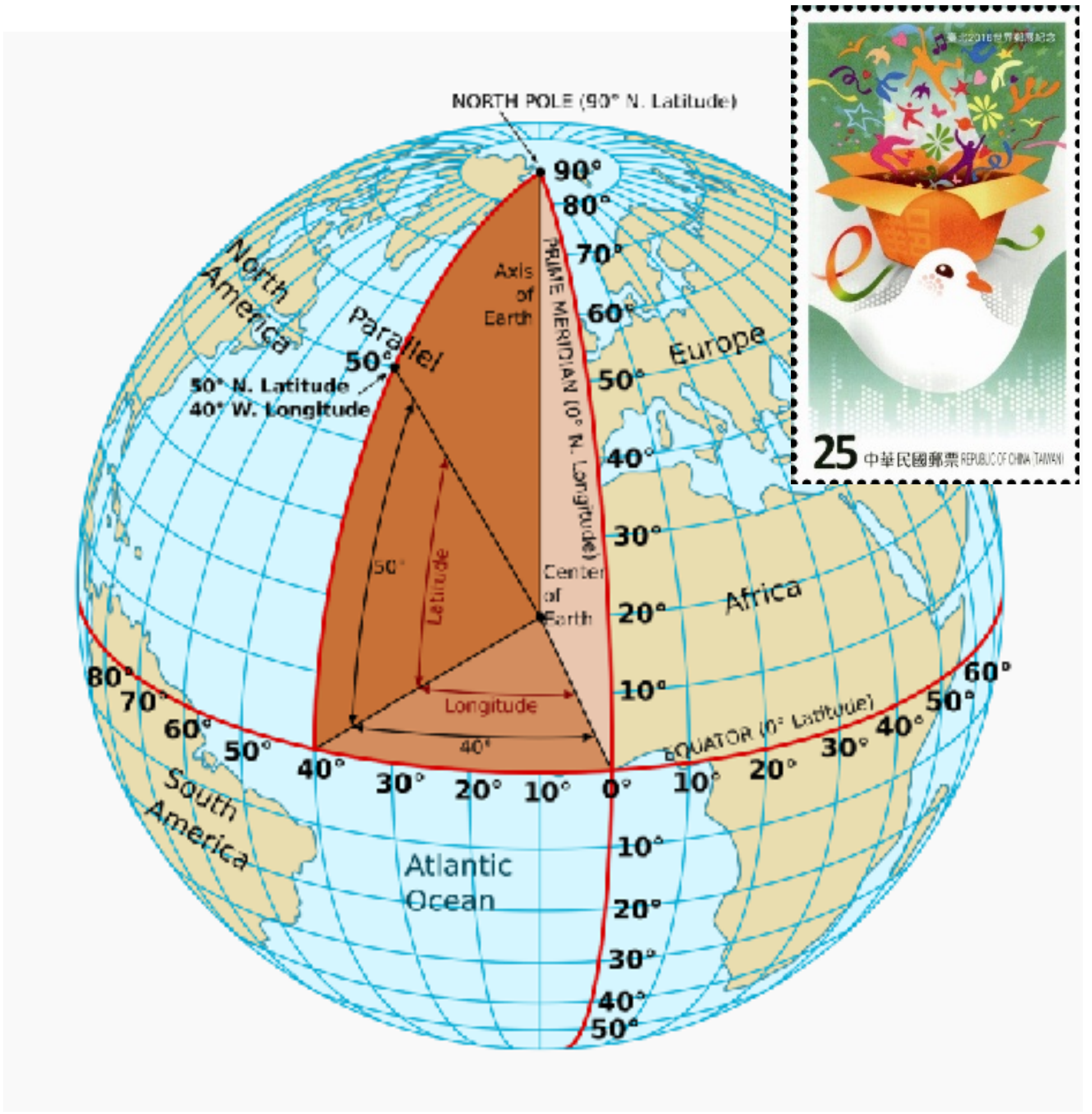
$$\alpha + \beta + \gamma = \pi + \frac{A}{R^2}$$

where A is the area of the triangle.

If A is very small or R is large, then $\alpha + \beta + \gamma \approx \pi$. This corresponds to drawing the triangle over a small part of the surface that is *locally* close to flat.

If $R \to \infty$, the sphere looks flat everywhere.





Main lesson: on a curved surface, the **radius of curvature** appears as an extra parameter in geometric formulae.

On a spherical surface, by definition, the radius of curvature is the same everywhere.

Curvature: $K \equiv 1/R^2$

$R > 0 \implies K > 0$: positive curvature.

$R = \infty \implies K = 0$: zero curvature.

Distance on a sphere

In 3-d spherical polar coordinates (r, θ, ϕ), the distance metric is

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$$

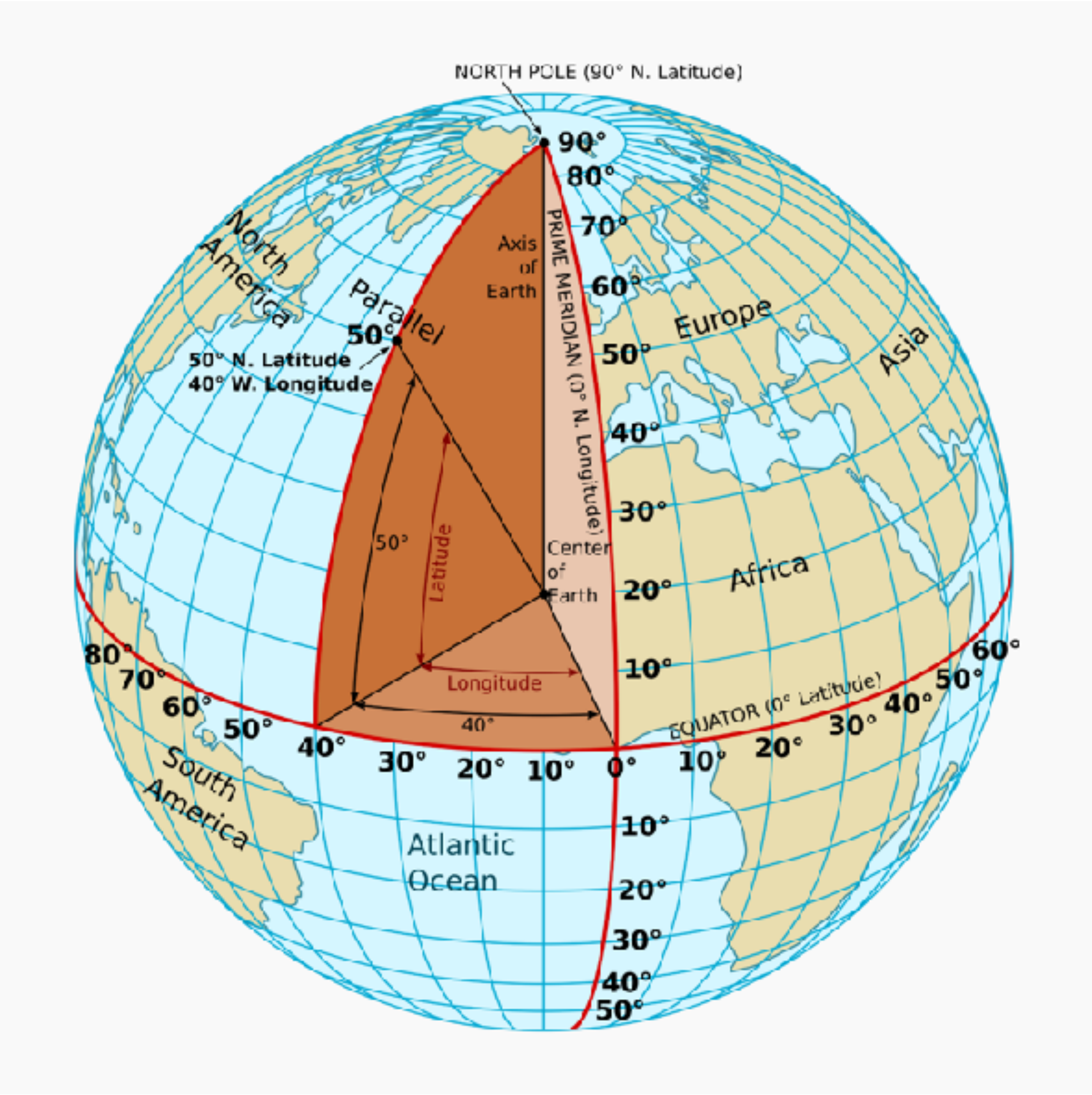
So on the **surface** of a sphere of radius R, the distance metric is

$$ds^2 = R^2\, d\theta^2 + R^2 \sin^2\theta\, d\phi^2$$

where (θ, ϕ) are angles that specify a location on the surface of the sphere (like latitude and longitude on Earth).

A journey around the axis of the Earth is longer at the equator than at the poles!

A spherical space has finite extent (is **closed**); a flat space has infinite extent (is **open**).



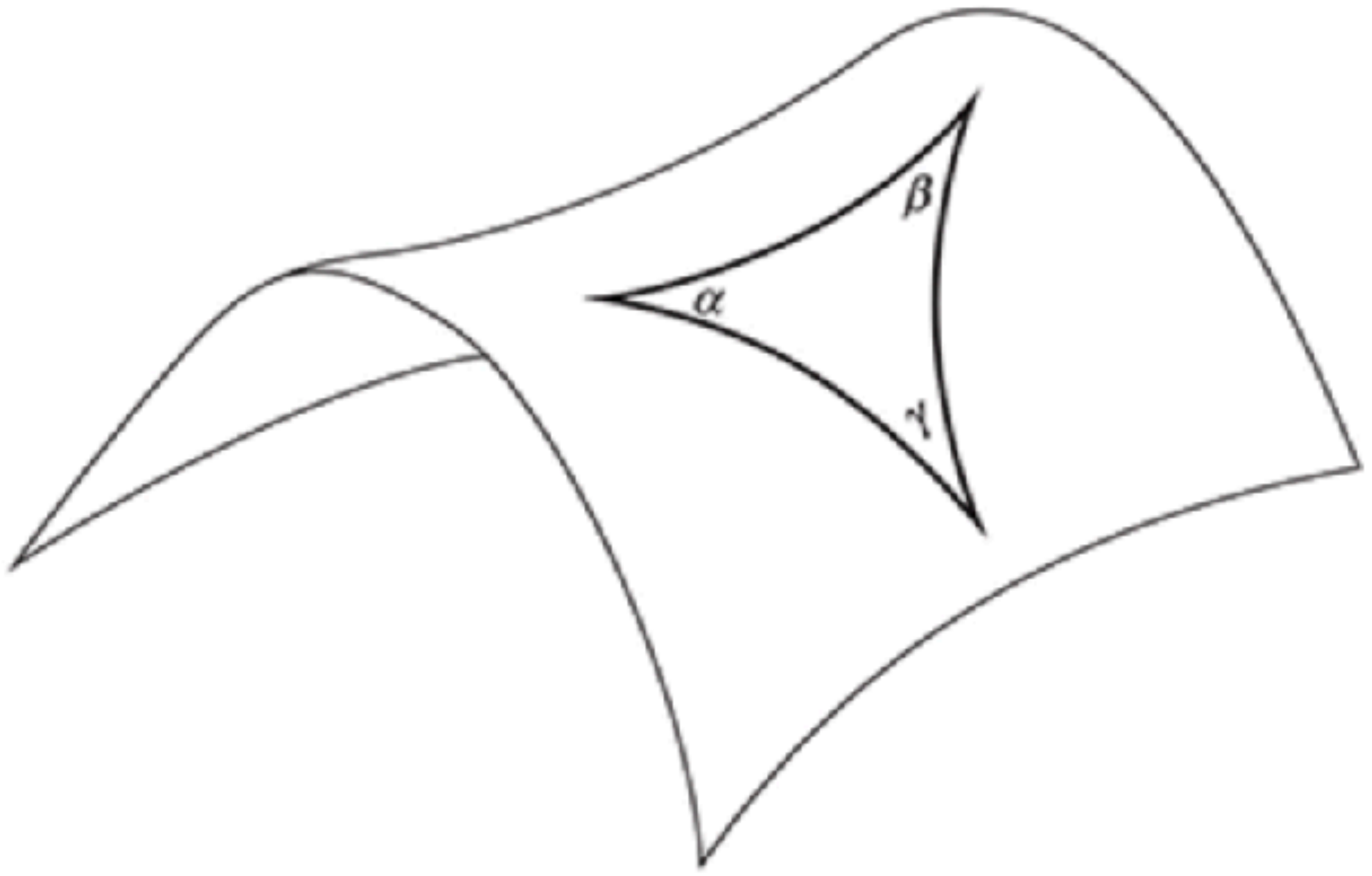
Positive and negative curvature

A sphere and a plane are two examples of surfaces with uniform curvature.

There is a third option: **negative** curvature, which corresponds to a **hyperbolic** surface.

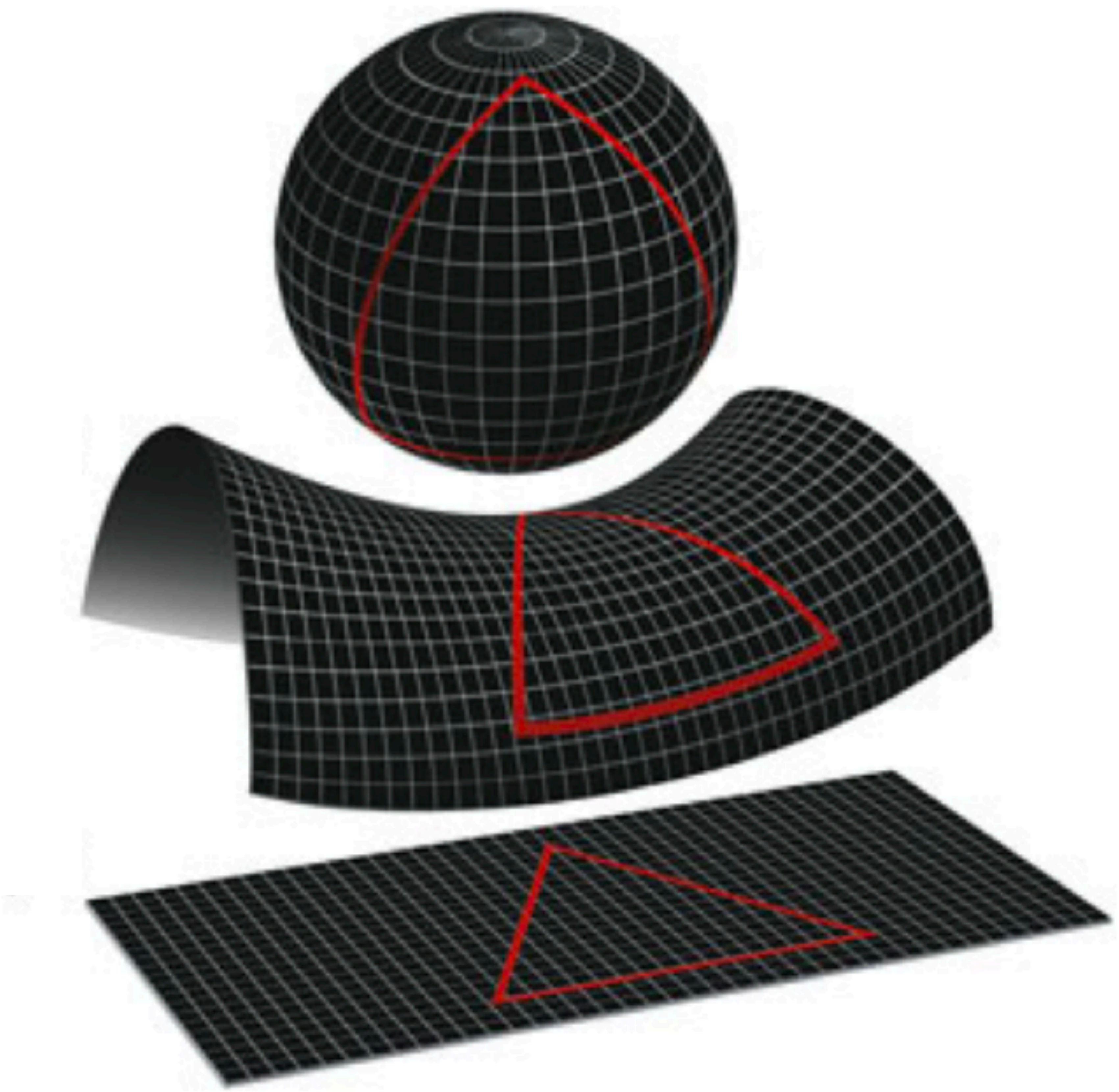
On such a surface, $\alpha + \beta + \gamma = \pi - \dfrac{A}{R^2}$

This equation treats R as real and positive; if we say the radius of curvature is imaginary, $R \to iR$, for a hyperbolic surface, then $K < 0$ and one formula $\alpha + \beta + \gamma = \pi + A/R^2$, works for all three cases.



A hyperbolic space has infinite extent (is **open**).

2-d curvature summary



Higher-dimensional curvature

The concept of curvature applies in more than 2 dimensions.

It is impossible to **draw** a 3-dimensional curved space!

To an observer living in such a space, however, the geometric signs curvature are the same: do the angles in a triangle add up to 180 degrees? Is the space **open** ($K \leq 0$) or **closed** ($K > 0$)?

Curvature as a property of a surface

The examples on the previous slides are 2-d curved surfaces embedded in a 3-dimensional space.

However, there is no need for a curved space to be **embedded** in a real higher dimensional space from which an observer could 'see' the curvature.

Specifically, the 3-d space of our Universe can be curved **regardless** of whether or not it is embedded in a 4-d (or higher) space. If it turns out our space has curvature (e.g. if angles of big triangles don't add up to π), this does not mean there is a 4^{th} space dimension.

(Sometimes it <u>might</u> be helpful to *pretend* there is a 4th space dimension when doing related maths — see e.g. the quick derivation of the FLRW metric from Weinberg repeated in Huterer's book)

We say that curvature (i.e. the sign and value of K; equivalently R) is an **intrinsic** property of space.

It is not possible to map a curved surface to a flat one (or a positively curved surface to a negatively curved one) without distorting angles and areas.

Metrics on curved 2-d surfaces

For uniform radius of curvature R:

Flat

$$ds^2 = dr^2 + r^2\, d\theta^2$$

Spherical

$$ds^2 = dr^2 + R^2 \sin^2(r/R)\, d\theta^2$$

Hyperbolic

$$ds^2 = dr^2 + R^2 \sinh^2(r/R)\, d\theta^2$$

Metrics on curved 3-d surfaces

For uniform radius of curvature R:

Flat

$$ds^2 = dr^2 + r^2[d\theta^2 + \sin^2\theta^2\, d\phi^2] = dr^2 + r^2\, d\Omega^2$$

Spherical

$$ds^2 = dr^2 + R^2 \sin^2(r/R)\, d\Omega^2$$

Hyperbolic

$$ds^2 = dr^2 + R^2 \sinh^2(r/R)\, d\Omega^2$$

Metrics on curved 3-d surfaces

We can write all the metrics on the previous slide as a single equation:

$$ds^2 = dr^2 + S_\kappa^2(r)\, d\Omega^2$$

by defining the function

$$S_\kappa(r) = \begin{cases} R \sin(r/R), & \text{positive curvature} \\ r, & \text{zero curvature} \\ R \sinh(r/R) & \text{negative curvature} \end{cases}$$

Often the sign of the curvature is represented with an integer $\kappa = \{-1, 0, 1\}$. In the above equation this is just a label, but it will show up again later.

General FLRW metric (version 1)

The general FLRW metric for a Universe with spatial curvature

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + S_\kappa^2(r)\, d\Omega^2]$$

$$S_\kappa(r) = \begin{cases} R_0 \sin(r/R_0), & \text{positive curvature} \\ r, & \text{zero curvature} \\ R_0 \sinh(r/R_0) & \text{negative curvature} \end{cases}$$

(note: $r \ll R_0 \implies S_\kappa \sim r$ regardless of κ)

$$R_0 \equiv R(t = t_0)$$

$$0 < a(t) < 1$$

General FLRW metric (version 2)

The same general FLRW metric for a Universe with spatial curvature can be written in a different-looking but mathematically identical way by writing $\varrho = S_k(r)$

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{d\varrho^2}{1-\kappa(\varrho/R_0)^2} + \varrho^2 d\Omega^2\right]$$

$$\kappa = \begin{cases} +1, & \text{positive curvature} \\ 0, & \text{zero curvature} \\ -1 & \text{negative curvature} \end{cases}$$

This version makes the "angular term" simpler at the expense of making the "radial term" more complicated. This is useful for certain problems.

Note that Eq. 2.19 in Huterer has a typo (change $\kappa \rightarrow k = \kappa/R_0^2$)

(https://public.websites.umich.edu/~huterer/BOOK/errata.txt)

Warning: notation!

Almost everyone uses slightly different notation when discussing the FLRW metric. Even worse, different books use the same small set of letters ($r, R, \chi, \theta, \kappa, k, K, \dots$) to mean physically different things!

r in Ryden (and our notes) $\rightarrow$ χ in Huterer
$S_k(r)$ in Ryden (and our notes) $\rightarrow$ r in Huterer

Sometimes the notation within one book is not even self-consistent!

Things are this way because:

1. *It doesn't matter for anything practical at the level we discuss — it's just notation;*

2. *<u>The difference for the flat case is trivial, and we mostly work with the flat case</u>.*

HOWEVER, <u>you</u> do not **need** to worry about this if you don't want to — unless you are trying **to read more than one book very carefully at the same time**, in which case it could be a big headache.

Summary

We live in an expanding Universe. Galaxies are (almost) comoving with the expansion and hence see other galaxies receding from them at a rate proportional to their separation.

The Hubble Parameter measures the rate of expansion at a given time: $H(t) = \dot{a}/a$. Present-day value $H_0 \simeq 68\,\mathrm{km\,s^{-1}\,Mpc}$ (from CMB) or .

An expanding (or contracting) curved Universe is a generic consequence of GR applied to a homogeneous, isotropic mass-energy distribution. Homogeneity implies uniform spatial curvature. The FLRW metric is the most generic form of metric for these conditions:

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + S_\kappa^2(r)\,d\Omega^2] \qquad ds^2 = -c^2dt^2 + a^2(t)\left[\frac{d\varrho^2}{1-\kappa(\varrho/R_0)^2} + \varrho^2 d\theta^2\right]$$

For a spatially flat Universe this reduces to $ds^2 = -c^2dt^2 + a^2(t)[dr^2 + r^2\,d\Omega^2]$.

We "factor out" the expansion using $a(t)$. We define **comoving coordinates** (at the present cosmic time, i.e. $a = 1$); we multiply these comoving coordinates by $a(t)$ when we want **physical coordinates** (separations) at a particular time.

Redshift is an observable consequence of expansion: $1 + z = \dfrac{1}{a(t)}$.

Next time

Once we specify a mass distribution, define a coordinate system and deduce the appropriate metric, we are ready to use the **GR field equations** to predict $a(t)$.

In the next lecture we will look at the Friedman equations, a recipe to construct $a(t)$ given different mixes of 'ingredients' for the mass-energy density.

For next time: Ch. 4 in Ryden, 2.6, 3.1, 3.2 in Huterer