8.286 Lecture 1
September 7, 2022

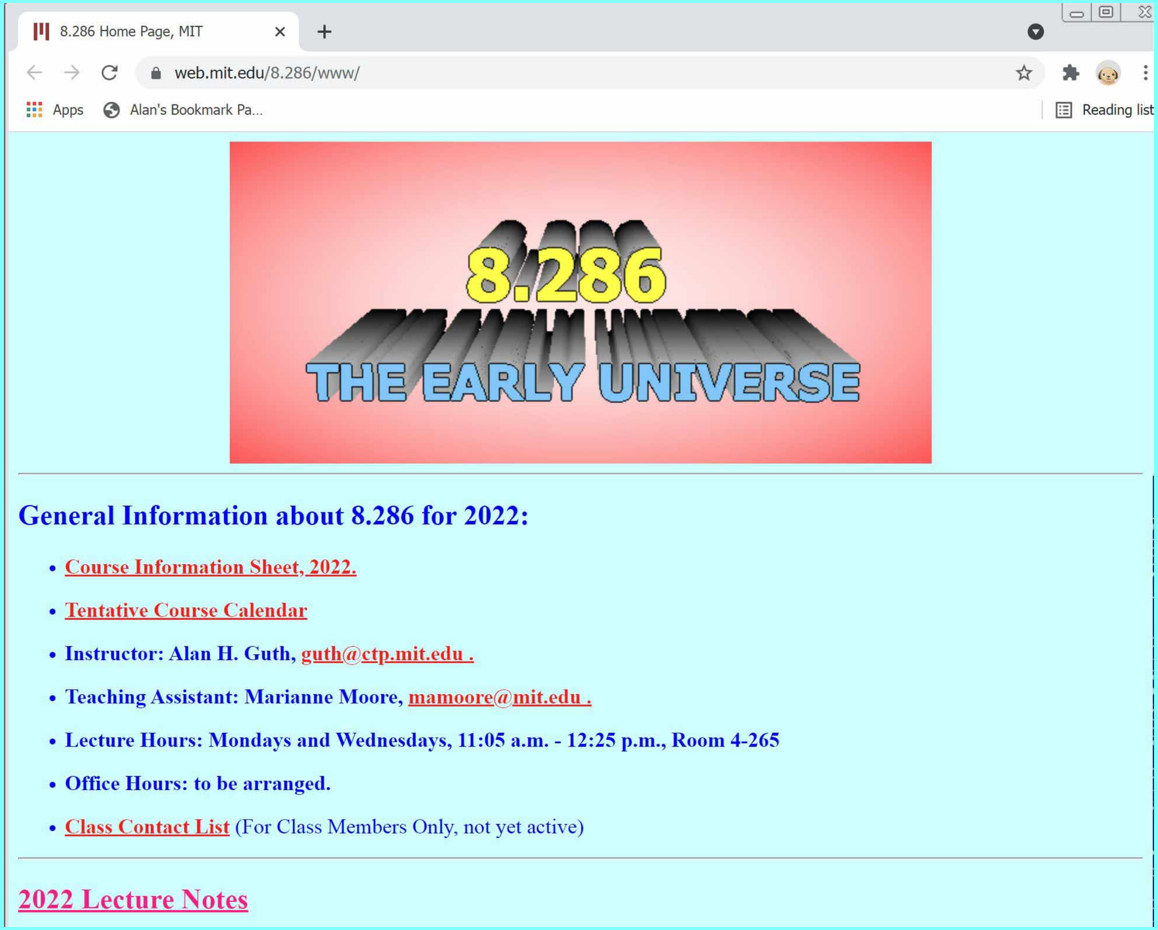
WELCOME TO 8.286!

OVERVIEW:

INFLATIONARY COSMOLOGY —
IS OUR UNIVERSE
PART OF A MULTIVERSE?

Course Website

https://web.mit.edu/8.286/www



Canvas

We will use Canvas for homework submission, and for showing your grades.

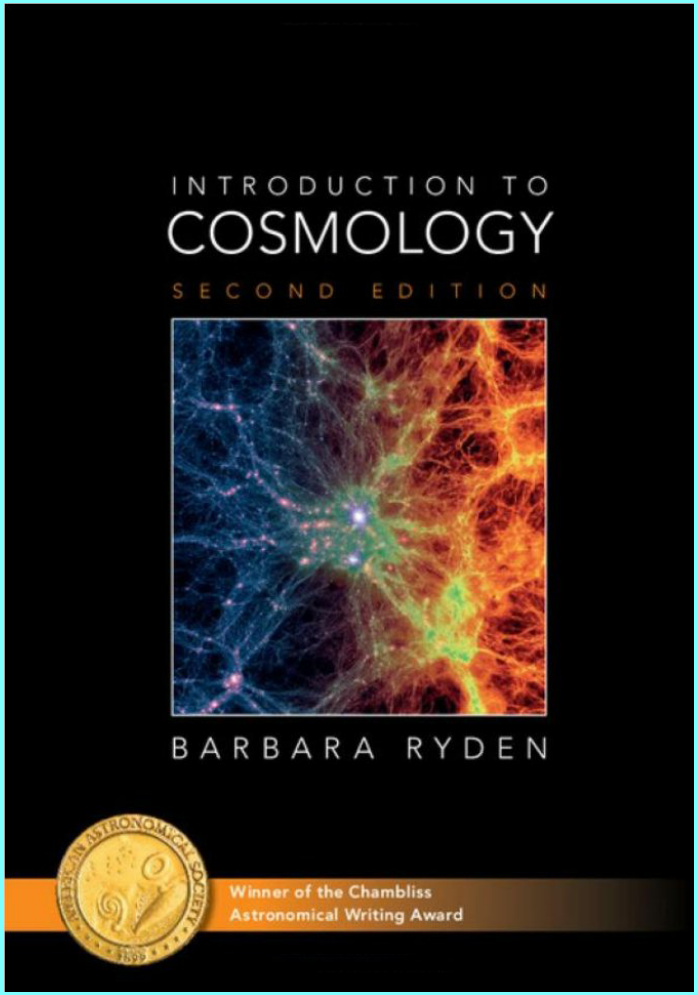
Course Staff

Lecturer: Me (Alan Guth), guth@ctp.mit.edu.

Teaching Assistant: Marianne Moore, mamoore@mit.edu.

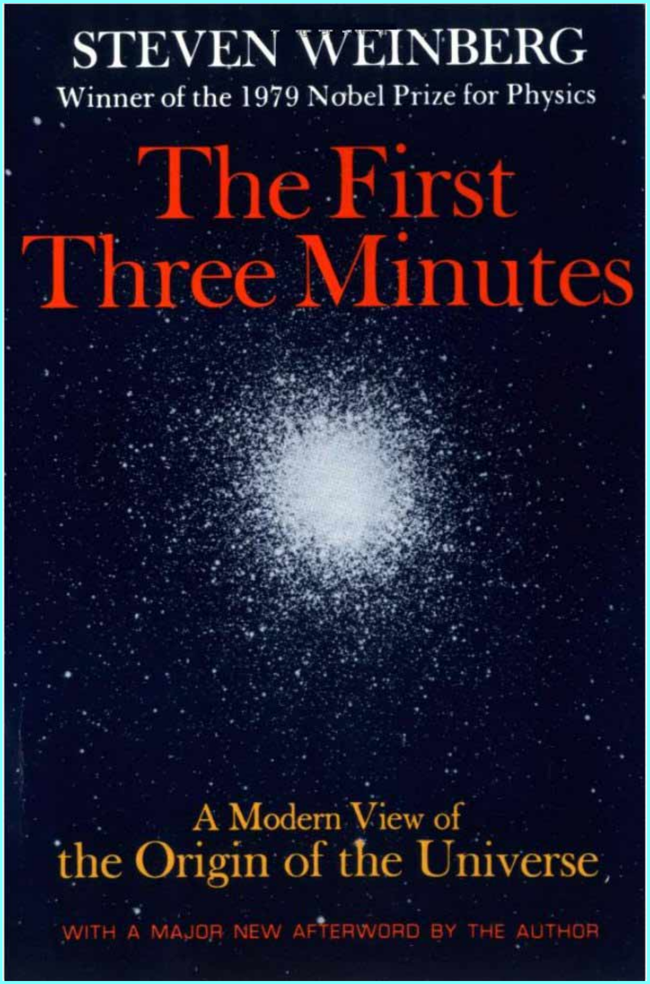
Required Textbook #1

Introduction to Cosmology, Second Edition (Cambridge University Press, 2016), by Barbara Ryden.



Required Textbook #2

The First Three Minutes, Second paperback edition, (Basic Books, 1993), by Steven Weinberg.



8.286 Lecture Notes

- Most of the course will be based on the 8.286 lecture notes, which will be posted as we go along. The first set is posted.
- All of the lecture notes from 2020 (and earlier years) are available on the website.
- Each quiz will have one question based on the reading in Ryden, Weinberg, and perhaps other brief sources near the end of the course. All of the readings from Weinberg will be fair game for such questions. You should think of Ryden's book as a supplement to the topics discussed in class, giving you a chance to see an alternative perspective. I will indicate some sections of Ryden's book that you are expected to read, and which will be fair game for the reading questions on quizzes.

Grading

Three "In-Class" Quizzes: 75%

The quizzes will tentatively be on the following dates:

Wednesday, October 5, 2022
Wednesday, November 9, 2022
Wednesday, December 7, 2022

If you have a problem of any kind with any of these dates, you should email me (Alan Guth) as soon as possible.

Problem Sets: 25%

No final exam.

Problem Sets

About 1 problem set per week, ten altogether, mostly due on Fridays at 5:00 pm Boston time.

There is a course calendar posted on the web page which shows the expected due dates for all the problem sets.

The first problem set is posted, and is due Friday, September 16.

The problem sets will not all be worth the same number of points. Your grade will be the total number of points you earn, compared to the maximum possible. That is, problem sets with more points will count a little more than the others.

All Problem Sets Will Count

Reason: Psets will be an integral part of the course, so you will miss something significant if you blow one of them off.

All Problem Sets Will Count

Reason: Psets will be an integral part of the course, so you will miss something significant if you blow one of them off.

However: I understand that you all have very overfilled lives. That's what MIT students are like! So, to make up for the fact that no problem set grades are being dropped, I will be VERY generous with extensions. If you are having an unusually busy week (or if you are sick, have a family crisis, or have a fight with your younger brother), just send me an email briefly describing the situation, and ask me for an extension.

If the solutions are posted before you turn in your problem set, you are on your honor not to look at the solutions, or discuss the problems with anyone who has (other than me or the other course staff), until you have turned it in.

Extra Credit

Some of the problem sets will offer additional problems for extra credit. We will keep track of the extra credit grades separately.

At the end of the course I will consult with Marianne Moore to set grade cuts based solely on the regular coursework. We will try to make sure that the grade cuts are fair with respect to this data set.

Then the extra credit grades will be added, allowing the grades to change upwards accordingly.

Finally, we will look at each student's grades individually, and we might decide to give a higher grade to some students who are slightly below a borderline. Students whose grades have improved significantly during the term, and students whose average has been pushed down by single low grade, and students who have been affected by adverse personal or medical problems will be the ones most likely to be boosted.

Homework Policy

I regard the problem sets primarily as an educational experience, rather than a mechanism of evaluation.

Your are encouraged — even strongly encouraged — to work on the homework in groups. I will be setting up a Class Contact webpage to help you find each other. But, you are each expected to **write up** your own solutions, even if you found those solutions as a group project.

We will ask you to indicate at the start of your homework the names of students whom you worked with, and students from outside the class whom you may have received help from. In that way the grader will know when to expect similarities in student's solutions.

8.286 Problem Solutions from previous years are strictly off limits, but other sources — textbooks, webpages — are okay, as long as you rewrite the solution in your own words.

A homework problem that appears to be copied from another student, from a previous year's solution, or copied from some other source without rewording might be given zero credit. Except in blatant cases, the first time you will be given a chance to redo it.

Remember that this homework policy does not necessarily apply to other classes.

Expect Two Emails

1) **Office Hours:** Marianne and I will each have one office hour a week, and the email questionnaire will be aimed at finding out what times work for you. I am using email because it allows more information transfer than Doodle or WhenToMeet. You will be able to tell us what hours do not work, what hours can work but are undesirable, and also what hours work best. If possible, we will arrange the pair of office hours so that all of you who respond will be able to attend at least one. There will also be a Piazza page where you can ask questions.

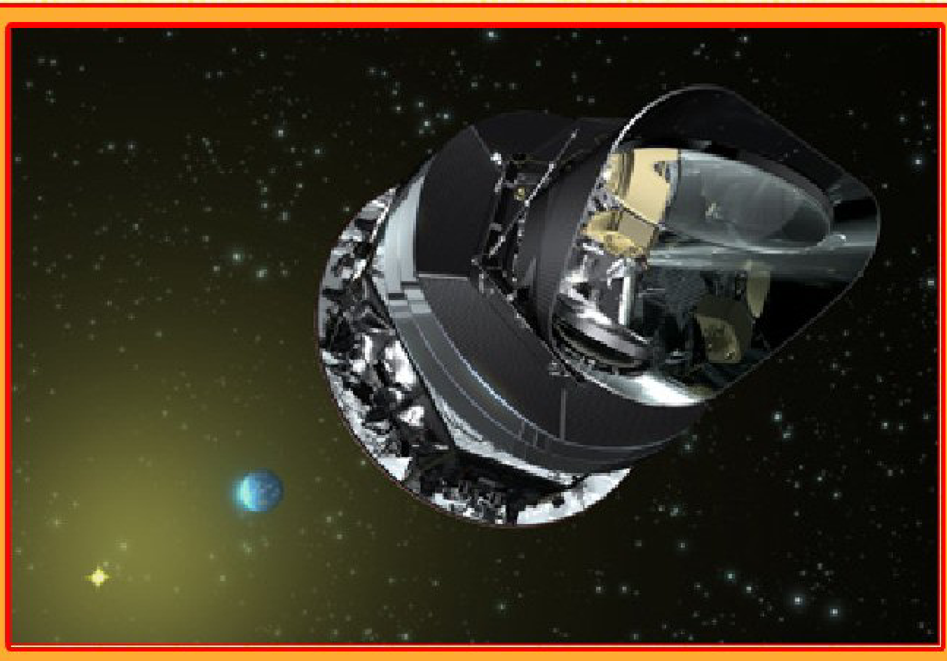
2) **Class Contact List:** This page, visible to all class members but no one else, is intended to help you find each other to discuss the course and to set up homework groups. It is Opt-In: if you don't want to be listed, you need not respond. The page will include photos (with your permission), contact information, and an optional comment describing what kind of interaction you are hoping to find.

☆ I will send these questionnaires to all registered students and listeners, and will post them on the website.

INFLATIONARY COSMOLOGY:

IS OUR UNIVERSE PART OF A MULTIVERSE?

PART 1



— Alan Guth —

Massachusetts Institute of Technology

8.286 Opening Lecture
September 7, 2022

The Standard Big Bang

What it is:

- Theory that the universe as we know it began 13-14 billion years ago. (Latest estimate: 13.80 ± 0.02 billion years, from the Planck satellite collaboration, 2018.)
- Initial state was a hot, dense, uniform soup of particles that filled space uniformly, and was expanding rapidly.

The Standard Big Bang

What it is:

- ☆ Theory that the universe as we know it began 13-14 billion years ago. (Latest estimate: 13.80 ± 0.02 billion years, from the Planck satellite collaboration, 2018.)
- ☆ Initial state was a hot, dense, uniform soup of particles that filled space uniformly, and was expanding rapidly.

What it describes:

- ☆ How the early universe expanded and cooled
- ☆ How the light chemical elements formed
- ☆ How the matter congealed to form stars, galaxies, and clusters of galaxies

What it doesn't describe:

- ☆ What caused the expansion? (The big bang theory describes only the **aftermath** of the bang.)
- ☆ Where did the matter come from? (The theory assumes that **all matter** existed from the very beginning.)

What it doesn't describe:

- ☆ What caused the expansion? (The big bang theory describes only the **aftermath** of the bang.)
- ☆ Where did the matter come from? (The theory assumes that **all matter** existed from the very beginning.)

In other words, it says nothing about what banged,

What it doesn't describe:

- ☆ What caused the expansion? (The big bang theory describes only the **aftermath** of the bang.)
- ☆ Where did the matter come from? (The theory assumes that **all matter** existed from the very beginning.)

In other words, it says nothing about what banged, why it banged,

What it doesn't describe:

- ☆ What caused the expansion? (The big bang theory describes only the **aftermath** of the bang.)
- ☆ Where did the matter come from? (The theory assumes that **all matter** existed from the very beginning.)

In other words, it says nothing about what banged, why it banged, or what happened before it banged!

What is Inflation?

☆ Inflation is a “prequel” to the conventional big bang picture.

What is Inflation?

- Inflation is a "prequel" to the conventional big bang picture.
- In particular, inflation is a theory about the bang of the big bang. That is, inflation is a possible answer to the question of what propelled the gigantic expansion of the big bang.

What is Inflation?

- ☆ Inflation is a "prequel" to the conventional big bang picture.
- ☆ In particular, inflation is a theory about the bang of the big bang. That is, inflation is a possible answer to the question of what propelled the gigantic expansion of the big bang.
- ☆ Inflation is also a possible answer to the question of where almost all of the matter came from.

What is Inflation?

- ☆ Inflation is a "prequel" to the conventional big bang picture.
- ☆ In particular, inflation is a theory about the bang of the big bang. That is, inflation is a possible answer to the question of what propelled the gigantic expansion of the big bang.
- ☆ Inflation is also a possible answer to the question of where almost all of the matter came from.
- ☆ Inflation is **NOT** a theory of the origin of the universe, but it can explain how the entire observed universe emerged from a patch only 10^{-28} cm across, with a mass of only a few grams.

What is Inflation?

- ☆ Inflation is a "prequel" to the conventional big bang picture.
- ☆ In particular, inflation is a theory about the bang of the big bang. That is, inflation is a possible answer to the question of what propelled the gigantic expansion of the big bang.
- ☆ Inflation is also a possible answer to the question of where almost all of the matter came from.
- ☆ Inflation is **NOT** a theory of the origin of the universe, but it can explain how the entire observed universe emerged from a patch only 10^{-28} cm across, with a mass of only a few grams.
- ☆ Inflation explains the expansion by means of

What is Inflation?

- ☆ Inflation is a "prequel" to the conventional big bang picture.
- ☆ In particular, inflation is a theory about the bang of the big bang. That is, inflation is a possible answer to the question of what propelled the gigantic expansion of the big bang.
- ☆ Inflation is also a possible answer to the question of where almost all of the matter came from.
- ☆ Inflation is **NOT** a theory of the origin of the universe, but it can explain how the entire observed universe emerged from a patch only 10^{-28} cm across, with a mass of only a few grams.
- ☆ Inflation explains the expansion by means of

Gravitational Repulsion.

Two Miracles of Physics

Two Miracles of Physics

Definition:

Two Miracles of Physics

Definition: A "miracle of physics" is a feature of the laws of physics which

Two Miracles of Physics

Definition: A "miracle of physics" is a feature of the laws of physics which

(a) was never taught to me when I was a student;

Two Miracles of Physics

Definition: A "miracle of physics" is a feature of the laws of physics which

(a) was never taught to me when I was a student; and

(b) is so far-reaching in its consequences that it can change our picture of the universe.

Miracle of Physics # 1: Gravitational Repulsion

- Since the advent of general relativity, physicists have known that gravity can act repulsively.
- In GR, pressures can create gravitational fields, and negative pressures create repulsive gravitational fields.
- Einstein used this possibility, in the form of the "cosmological constant," to build a static mathematical model of the universe, with repulsive gravity preventing its collapse.
- Modern particle physics suggests that at superhigh energies there should be many states with negative pressures, creating repulsive gravity.

☆ Inflation proposes that a patch of repulsive gravity material existed in the early universe — for inflation at the grand unified theory scale ($\sim 10^{16}$ GeV), the patch needs to be only as large as 10^{-28} cm. (Since any such patch is enlarged fantastically by inflation, the initial density or probability of such patches can be very low.)

- ☆ Inflation proposes that a patch of repulsive gravity material existed in the early universe — for inflation at the grand unified theory scale ($\sim 10^{16}$ GeV), the patch needs to be only as large as 10^{-28} cm. (Since any such patch is enlarged fantastically by inflation, the initial density or probability of such patches can be very low.)

  1 GeV $\approx$ mass energy of a proton.

- ☆ Inflation proposes that a patch of repulsive gravity material existed in the early universe — for inflation at the grand unified theory scale ($\sim 10^{16}$ GeV), the patch needs to be only as large as 10^{-28} cm. (Since any such patch is enlarged fantastically by inflation, the initial density or probability of such patches can be very low.)

  1 GeV $\approx$ mass energy of a proton.

- ☆ The gravitational repulsion created by this material was the driving force behind the big bang. The repulsion drove it into exponential expansion, doubling in size every 10^{-37} second or so!

- ☆ The patch expanded exponentially by a factor of at least 10^{28} ($\sim$ 100 doublings), but it could have expanded much more. Inflation lasted maybe 10^{-35} second, and at the end, the region destined to become the presently observed universe was about the size of a marble.

- ☆ The repulsive-gravity material is unstable, so it decayed like a radioactive substance, ending inflation. The decay released energy which produced ordinary particles, forming a hot, dense "primordial soup." Standard cosmology began.

- ☆ The patch expanded exponentially by a factor of at least 10^{28} ($\sim$ 100 doublings), but it could have expanded much more. Inflation lasted maybe 10^{-35} second, and at the end, the region destined to become the presently observed universe was about the size of a marble.

- ☆ The repulsive-gravity material is unstable, so it decayed like a radioactive substance, ending inflation. The decay released energy which produced ordinary particles, forming a hot, dense "primordial soup." Standard cosmology began.

  Caveat: The decay happens almost everywhere, but not everywhere — we will come back to this subtlety, which is the origin of eternal inflation.

★ After the repulsive-gravity material decayed, the universe continued to coast and cool from then onward.

- After the repulsive-gravity material decayed, the universe continued to coast and cool from then onward.
- Key feature: During the exponential expansion, the density of matter and energy did NOT thin out. The density of the repulsive gravity material was **not lowered** as it expanded!

- After the repulsive-gravity material decayed, the universe continued to coast and cool from then onward.
- Key feature: During the exponential expansion, the density of matter and energy did NOT thin out. The density of the repulsive gravity material was **not lowered** as it expanded!
- Although more and more mass/energy appeared as the repulsive-gravity material expanded, total energy was conserved!

- After the repulsive-gravity material decayed, the universe continued to coast and cool from then onward.
- Key feature: During the exponential expansion, the density of matter and energy did NOT thin out. The density of the repulsive gravity material was **not lowered** as it expanded!
- Although more and more mass/energy appeared as the repulsive-gravity material expanded, total energy was conserved! HOW????

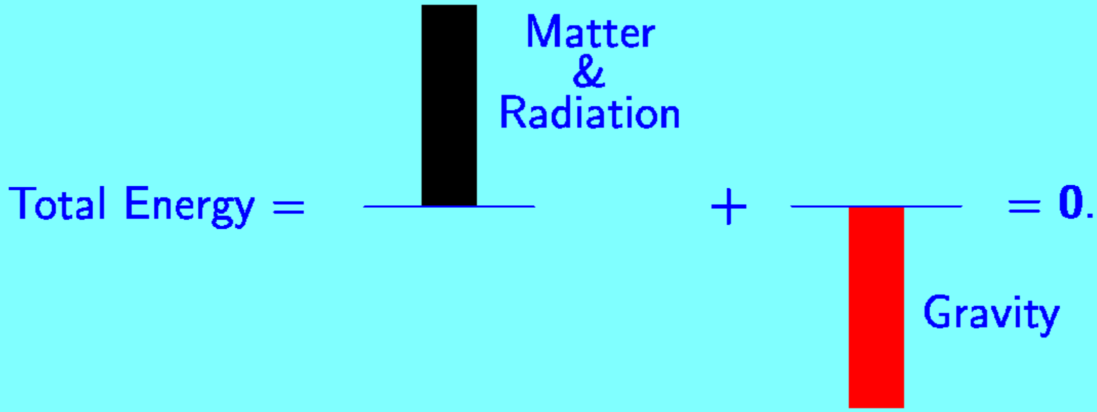
Miracle of Physics #2: Energy is Conserved, But Not Always Positive

Miracle of Physics #2: Energy is Conserved, But Not Always Positive

- The energy of a gravitational field is negative (both in Newtonian gravity and in general relativity).
- The negative energy of gravity canceled the positive energy of matter, so the total energy was constant and possibly zero.

Miracle of Physics #2: Energy is Conserved, But Not Always Positive

- ☆ The energy of a gravitational field is negative (both in Newtonian gravity and in general relativity).
- ☆ The negative energy of gravity canceled the positive energy of matter, so the total energy was constant and possibly zero.
- ☆ The total energy of the universe today is consistent with zero. Schematically,



Evidence for Inflation

1) **Large scale uniformity.** The cosmic background radiation is uniform in temperature to one part in 100,000. It was released when the universe was about 400,000 years old. In standard cosmology without inflation, a mechanism to establish this uniformity would need to transmit energy and information at about 100 times the speed of light.

Evidence for Inflation

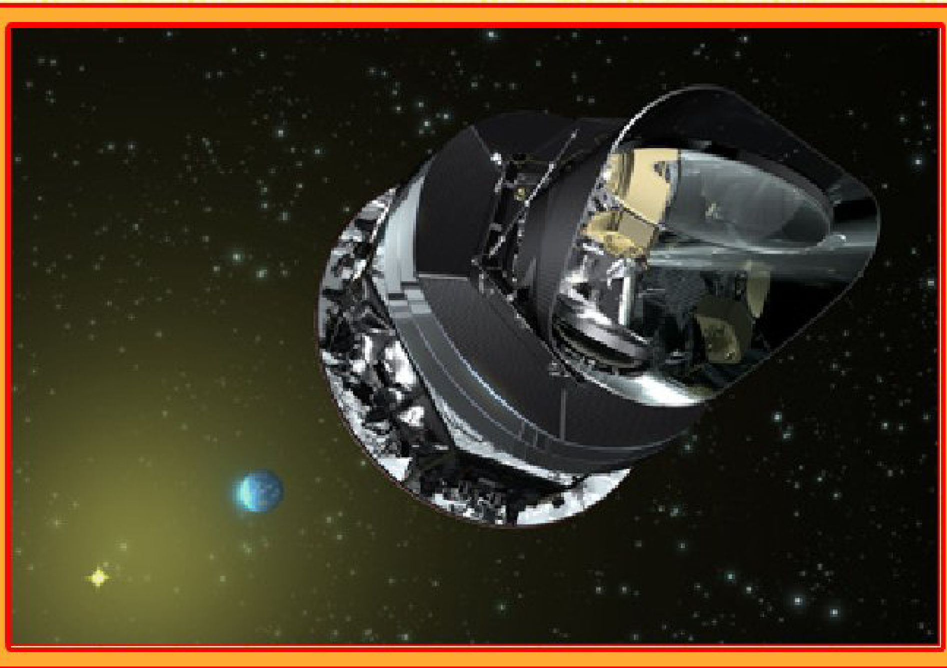
1) **Large scale uniformity.** The cosmic background radiation is uniform in temperature to one part in 100,000. It was released when the universe was about 400,000 years old. In standard cosmology without inflation, a mechanism to establish this uniformity would need to transmit energy and information at about 100 times the speed of light.

Inflationary Solution: In inflationary models, the universe begins so small that uniformity is easily established — just like the air in the lecture hall spreading to fill it uniformly. Then inflation stretches the region to be large enough to include the visible universe.

INFLATIONARY COSMOLOGY:

IS OUR UNIVERSE PART OF A MULTIVERSE?

PART 2



— Alan Guth —

Massachusetts Institute of Technology

8.286 Lecture 2
September 12, 2022

SUMMARY OF LAST LECTURE

The Conventional Big Bang Theory (i.e., without inflation): Really describes only the aftermath of a bang: It says nothing about what banged, why it banged, or what happened before it banged. The description begins with a hot dense uniform soup of particles filling an expanding space.

Cosmic Inflation: The prequel, describes how **repulsive gravity** — a consequence of **negative pressure** — could have driven a tiny patch of the early universe into exponential expansion. The total energy would be very small or maybe zero, with the **negative energy of the cosmic gravitational field** canceling the energy of matter.

Evidence for Inflation

1) **Large scale uniformity.** The cosmic background radiation is uniform in temperature to one part in 100,000. It was released when the universe was about 400,000 years old. In standard cosmology without inflation, a mechanism to establish this uniformity would need to transmit energy and information at about 100 times the speed of light.

Evidence for Inflation

1) **Large scale uniformity.** The cosmic background radiation is uniform in temperature to one part in 100,000. It was released when the universe was about 400,000 years old. In standard cosmology without inflation, a mechanism to establish this uniformity would need to transmit energy and information at about 100 times the speed of light.

Inflationary Solution: In inflationary models, the universe begins so small that uniformity is easily established — just like the air in the lecture hall spreading to fill it uniformly. Then inflation stretches the region to be large enough to include the visible universe.

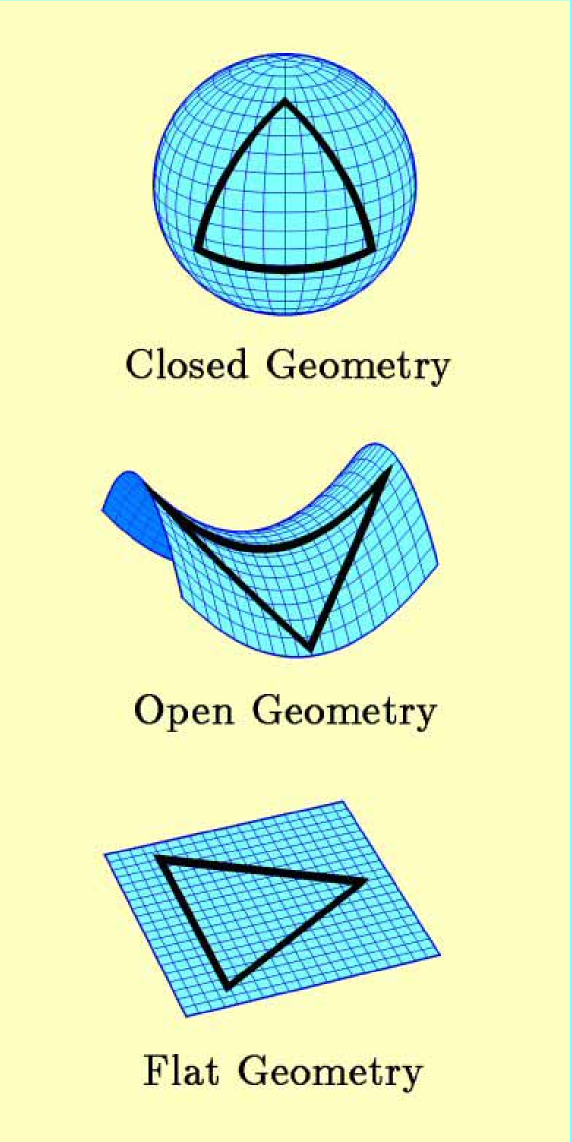
2) **“Flatness problem:”**

Why was the early universe so **FLAT?**

2) **“Flatness problem:”**

Why was the early universe so $\boxed{\textbf{FLAT?}}$

- ☆ If we assume that the universe is homogeneous (same in all places) and isotropic (same in all directions), then there are only three possible geometries: closed, open, or flat.



2) **"Flatness problem:"**

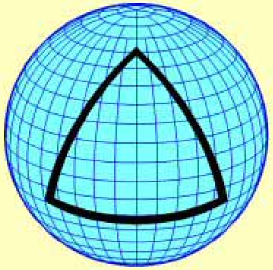
Why was the early universe so **FLAT?**

- ☆ If we assume that the universe is homogeneous (same in all places) and isotropic (same in all directions), then there are only three possible geometries: closed, open, or flat.

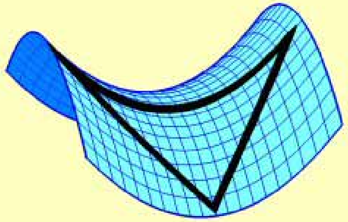
- ☆ According to general relativity, the flatness of the universe is related to its mass density:

$$\Omega(Omega) = \frac{\text{actual mass density}}{\text{critical mass density}} \, ,$$

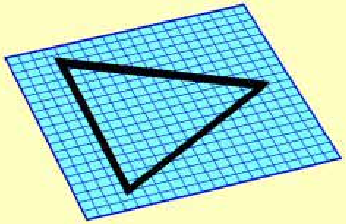
where the "critical density" depends on the expansion rate. $\Omega = 1$ is flat, $\Omega > 1$ is closed, $\Omega < 1$ is open.



Closed Geometry

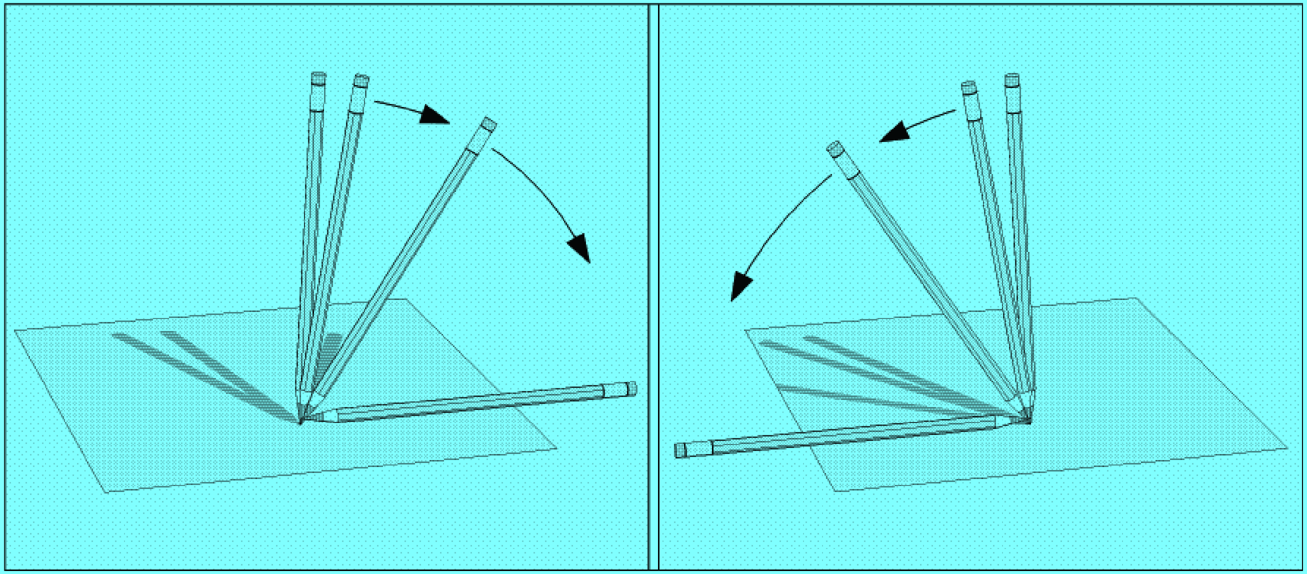


Open Geometry



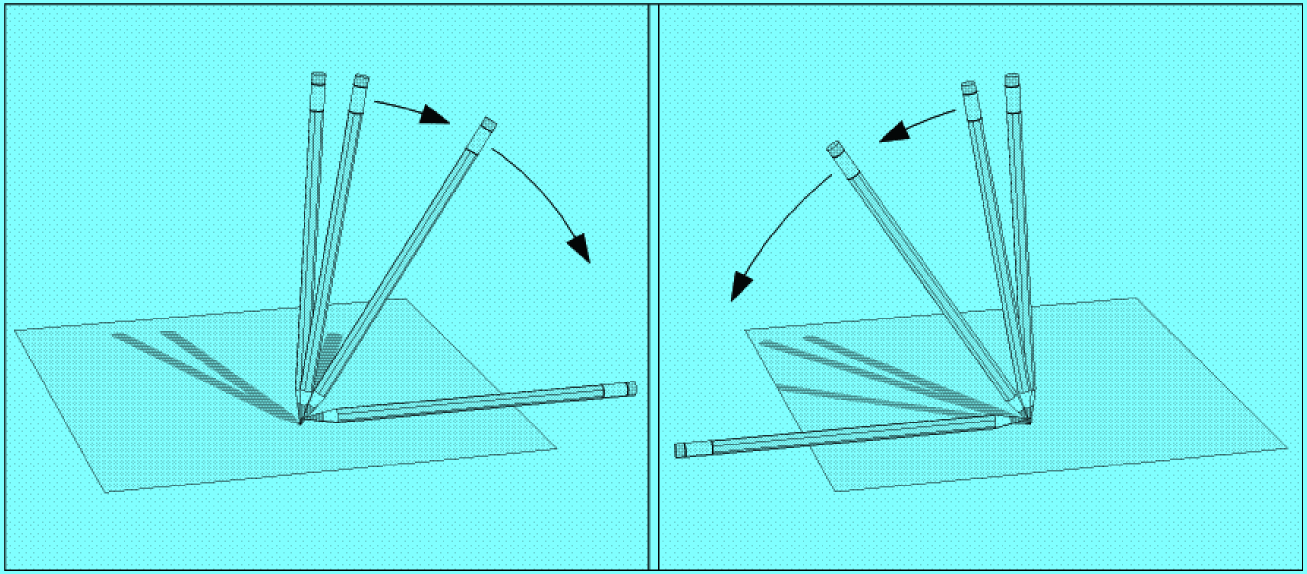
Flat Geometry

- A universe at the critical density is like a pencil balancing on its tip:



- If Ω in the early universe was slightly below 1, it would rapidly fall to zero — and no galaxies would form.
- If Ω was slightly greater than 1, it would rapidly rise to infinity, the universe would recollapse, and no galaxies would form.

- A universe at the critical density is like a pencil balancing on its tip:



- If Ω in the early universe was slightly below 1, it would rapidly fall to zero — and no galaxies would form.
- If Ω was slightly greater than 1, it would rapidly rise to infinity, the universe would recollapse, and no galaxies would form.
- To be even within a factor of 10 of the critical density today (which is what we knew in 1980), at one second after the big bang, Ω must have been equal to one **to 15 decimal places!**

☆ $\boxed{\textbf{Inflationary Solution:}}$ Since inflation makes gravity become repulsive, the evolution of Ω changes, too. Ω is driven towards one, extremely rapidly. It could begin at almost any value.

- **Inflationary Solution:** Since inflation makes gravity become repulsive, the evolution of Ω changes, too. Ω is driven towards one, extremely rapidly. It could begin at almost any value.
- Since the mechanism by which inflation explains the flatness of the early universe almost always overshoots, it predicts that even today the universe should have a critical density.

- **Inflationary Solution:** Since inflation makes gravity become repulsive, the evolution of Ω changes, too. Ω is driven towards one, extremely rapidly. It could begin at almost any value.
- Since the mechanism by which inflation explains the flatness of the early universe almost always overshoots, it predicts that even today the universe should have a critical density.
- Until 1998, observation pointed to $\Omega \approx 0.2$–0.3.

☆ **Inflationary Solution:** Since inflation makes gravity become repulsive, the evolution of Ω changes, too. Ω is driven towards one, extremely rapidly. It could begin at almost any value.

☆ Since the mechanism by which inflation explains the flatness of the early universe almost always overshoots, it predicts that even today the universe should have a critical density.

☆ Until 1998, observation pointed to $\Omega \approx 0.2$–0.3.

☆ Latest observation by Planck satellite (2018, combined with other astronomical observations):

$$\boxed{\Omega = 0.9993 \pm 0.0037 \text{ (95\%confidence)}}$$

- **Inflationary Solution:** Since inflation makes gravity become repulsive, the evolution of Ω changes, too. Ω is driven towards one, extremely rapidly. It could begin at almost any value.
- Since the mechanism by which inflation explains the flatness of the early universe almost always overshoots, it predicts that even today the universe should have a critical density.
- Until 1998, observation pointed to $\Omega \approx 0.2$–0.3.
- Latest observation by Planck satellite (2018, combined with other astronomical observations):

$$\boxed{\Omega = 0.9993 \pm 0.0037 \text{ (95\%confidence)}}$$

- New ingredient: Dark Energy. In 1998 it was discovered that the expansion of the universe has been accelerating for about the last 5 billion years. The “Dark Energy” is the energy causing this to happen.

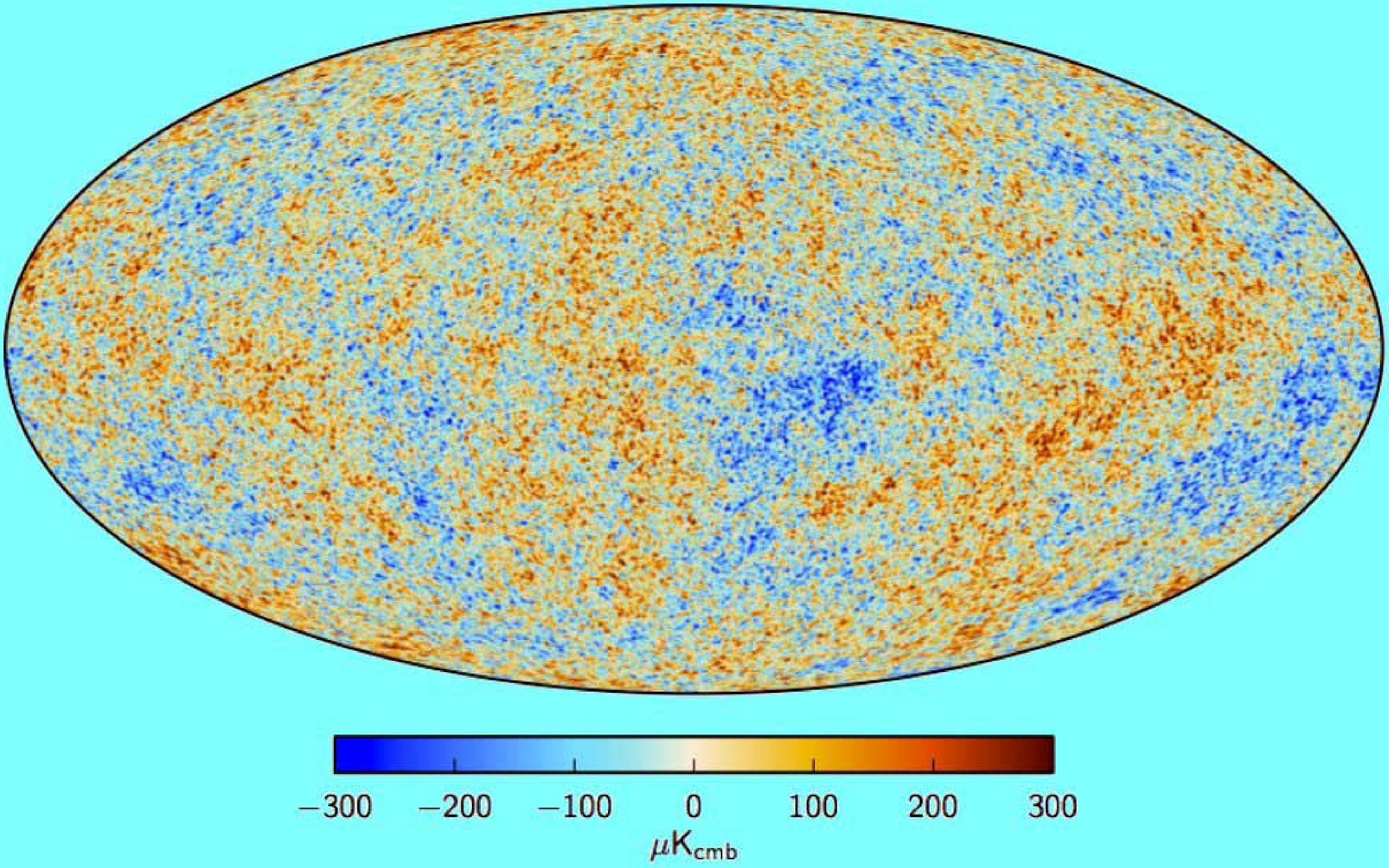
3) **Small scale nonuniformity:** Can be measured in the cosmic background radiation. The intensity is almost uniform across the sky, but there are small ripples. Although these ripples are only at the level of 1 part in 100,000, these nonuniformities are now detectable! Where do they come from?

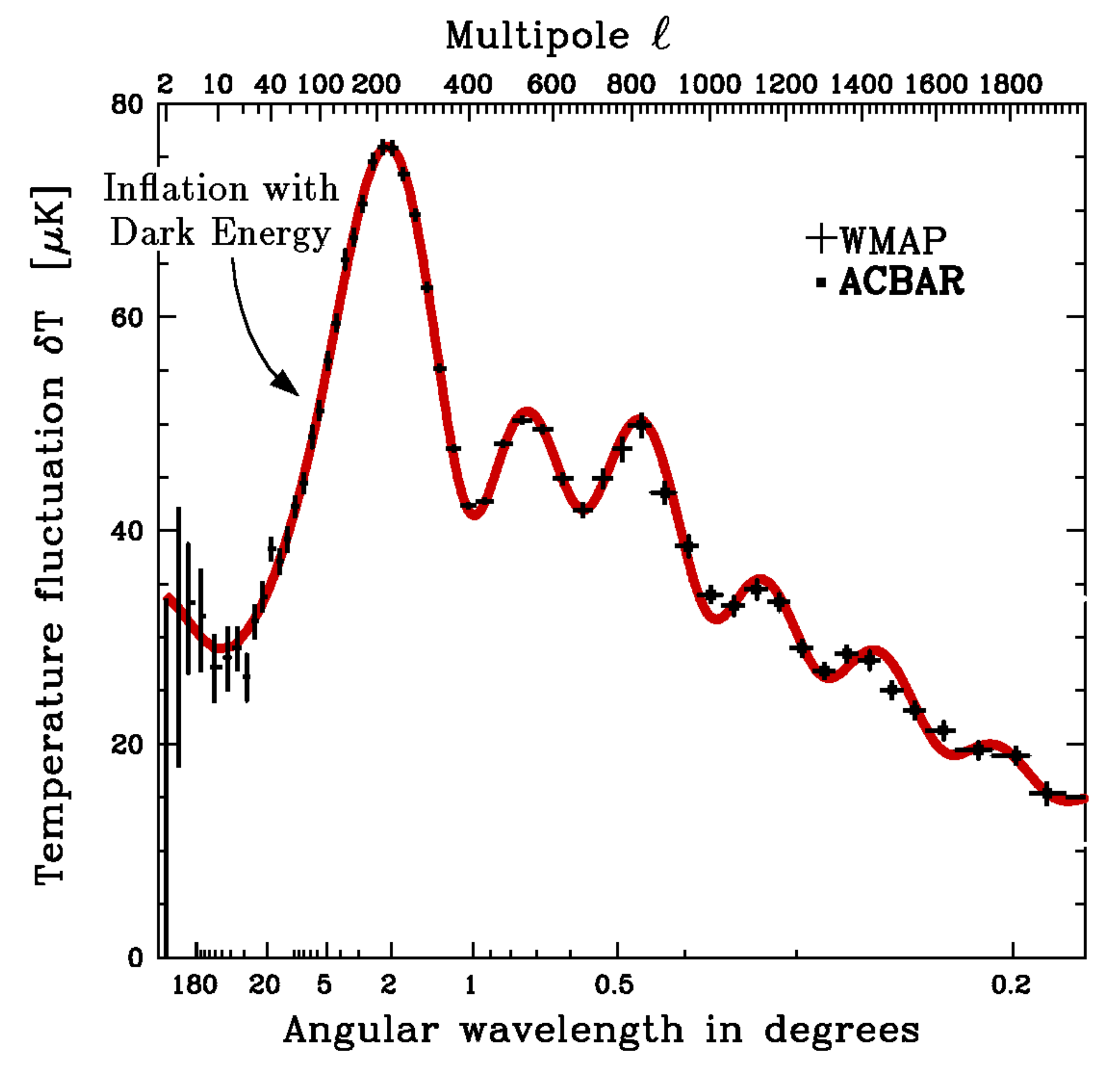
3) **Small scale nonuniformity:** Can be measured in the cosmic background radiation. The intensity is almost uniform across the sky, but there are small ripples. Although these ripples are only at the level of 1 part in 100,000, these nonuniformities are now detectable! Where do they come from?

Inflationary Solution: Inflation attributes these ripples to *quantum fluctuations.* Inflation makes generic predictions for the spectrum of these ripples (i.e., how the intensity varies with wavelength). The data measured so far agree beautifully with inflation.

Ripples in the Cosmic Microwave Background

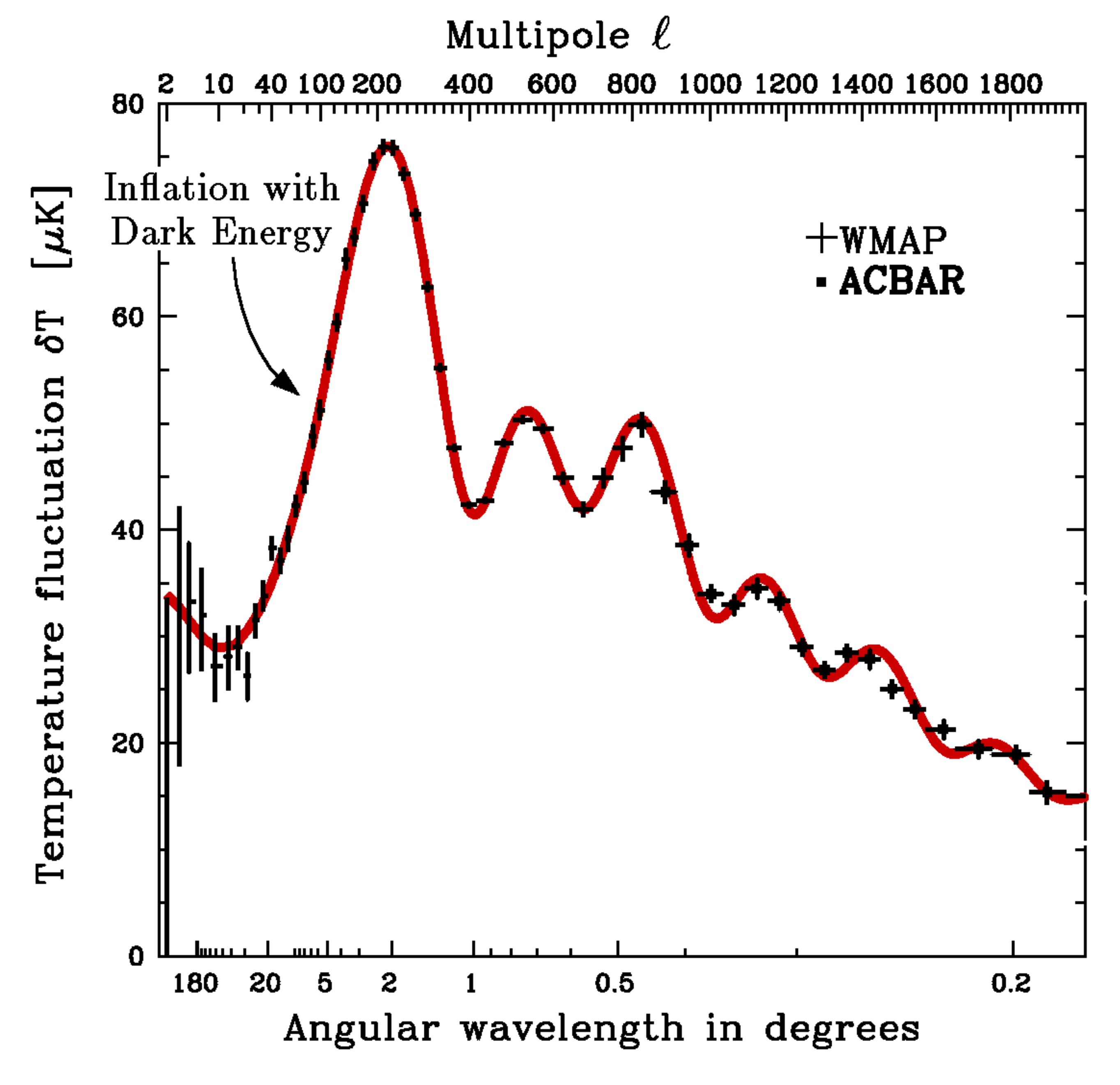
Planck Collaboration: The *Planck* mission





CMB: Comparison of Theory and Experiment

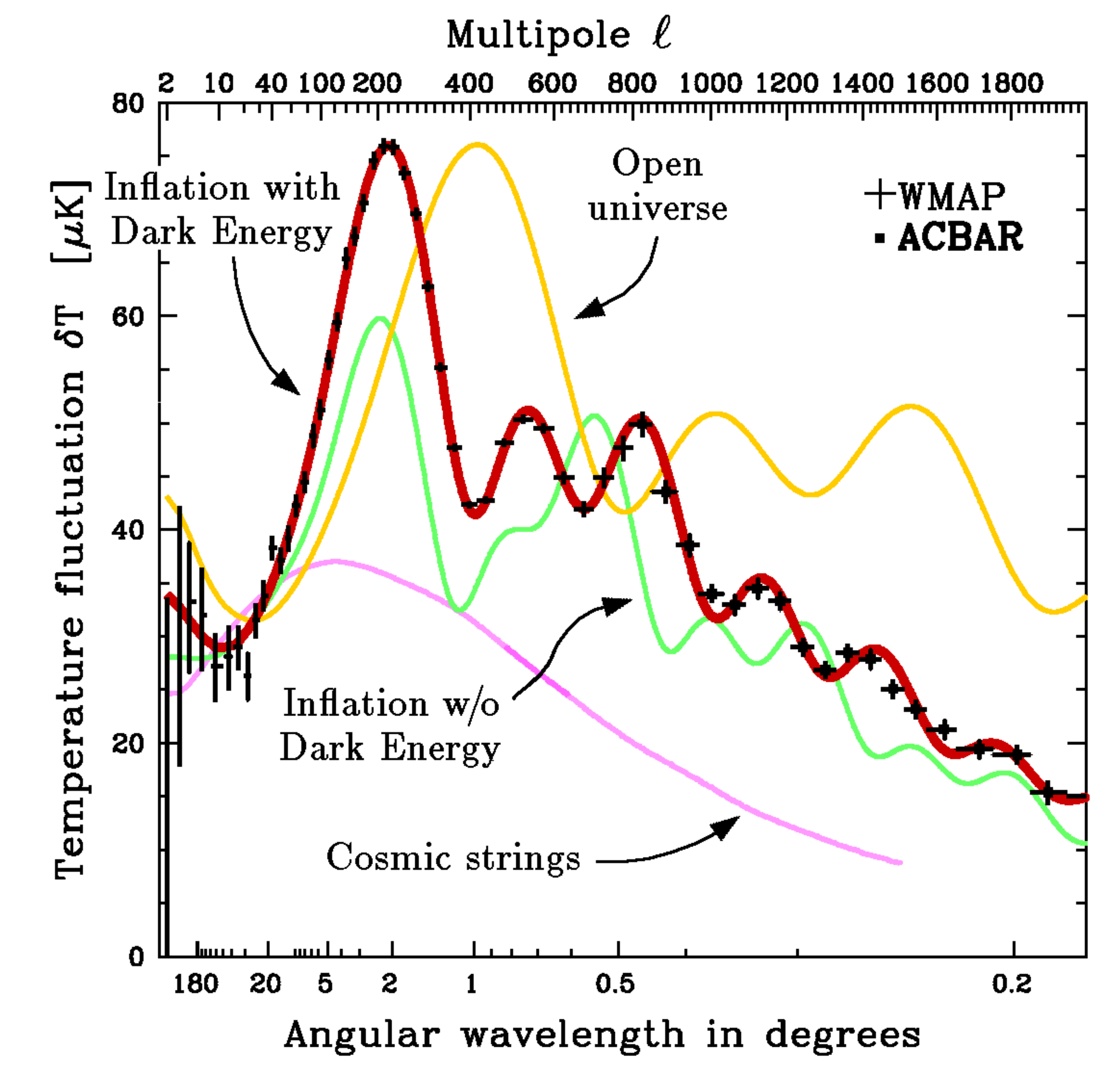
Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data.



CMB: Comparison of Theory and Experiment

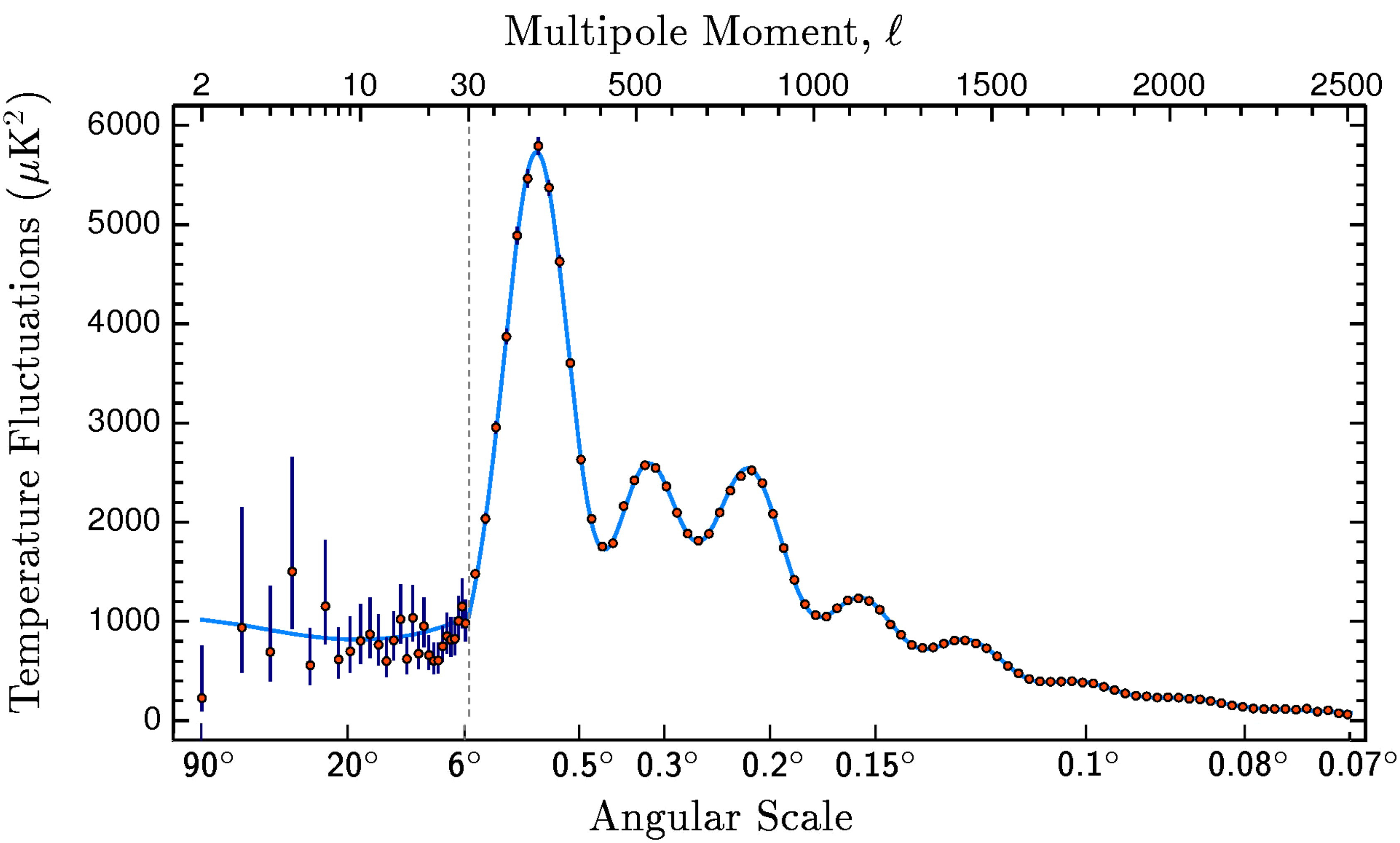
Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data.

CMB: Comparison of Theory and Experiment



Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data.

Spectrum of CMB Ripples



Planck Collaboration, 2018

Gravitational Waves:

Gravitational Waves: Came

Gravitational Waves: Came and Went

Gravitational Waves: Came and Went

March 17, 2014: The BICEP2 press conference announced the detection of swirly patterns (B-modes) in the polarization of the CMB, indicating gravitational waves from the early universe, in agreement with inflation.

Gravitational Waves: Came and Went

March 17, 2014: The BICEP2 press conference announced the detection of swirly patterns (B-modes) in the polarization of the CMB, indicating gravitational waves from the early universe, in agreement with inflation.

Result: After accounting for their best estimate of contamination due to dust, they found a tensor/scalar ratio $r = 0.16^{+0.06}_{-0.05}$, with $r = 0$ disfavored at 5.9σ.

Gravitational Waves: Came and Went

March 17, 2014: The BICEP2 press conference announced the detection of swirly patterns (B-modes) in the polarization of the CMB, indicating gravitational waves from the early universe, in agreement with inflation.

Result: After accounting for their best estimate of contamination due to dust, they found a tensor/scalar ratio $r = 0.16^{+0.06}_{-0.05}$, with $r = 0$ disfavored at 5.9σ.

Translation: the probability that there are no primordial gravitational waves is only about one in a billion.

Gravitational Waves: Came and Went

March 17, 2014: The BICEP2 press conference announced the detection of swirly patterns (B-modes) in the polarization of the CMB, indicating gravitational waves from the early universe, in agreement with inflation.

Result: After accounting for their best estimate of contamination due to dust, they found a tensor/scalar ratio $r = 0.16^{+0.06}_{-0.05}$, with $r = 0$ disfavored at 5.9σ.

Translation: the probability that there are no primordial gravitational waves is only about one in a billion.

April 14, 2015: *A Joint Analysis of BICEP2/Keck Array and Planck Data*: "We find strong evidence for dust and no statistically significant evidence for tensor modes."

The search for B-modes is still on, and if they are found they will provide additional strong evidence for inflation, and a tool for probing the details of inflation.

The search for B-modes is still on, and if they are found they will provide additional strong evidence for inflation, and a tool for probing the details of inflation.

Current limit: $r < 0.036$.

The search for B-modes is still on, and if they are found they will provide additional strong evidence for inflation, and a tool for probing the details of inflation.

Current limit: $r < 0.036$. Future sensitivity: if $r > 0.001$, it can be found by about 2024 (CMB Stage 4).

The search for B-modes is still on, and if they are found they will provide additional strong evidence for inflation, and a tool for probing the details of inflation.

Current limit: $r < 0.036$. Future sensitivity: if $r > 0.001$, it can be found by about 2024 (CMB Stage 4).

If B-modes are not found, that is not evidence against inflation: many inflationary models predict a B-mode intensity much smaller than 0.001. In 2018 I was involved in a paper about an inflationary model that gave $r \sim 10^{-29}$!

Inflation Suggests a Multiverse

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Roughly speaking, inflation is driven by a metastable state, which decays with some half-life.

After one half-life, half of the inflating material has become normal, noninflating matter, but the half that remains has continued to expand exponentially. It is vastly larger than it was at the beginning.

Once started, the inflation goes on FOREVER, with pieces of the inflating region breaking off and producing "pocket universes."

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Roughly speaking, inflation is driven by a metastable state, which decays with some half-life.

After one half-life, half of the inflating material has become normal, noninflating matter, but the half that remains has continued to expand exponentially. It is vastly larger than it was at the beginning.

Once started, the inflation goes on FOREVER, with pieces of the inflating region breaking off and producing "pocket universes."

We would be living in one of the infinity of pocket universes.

The Cosmological Constant Problem

- In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.

The Cosmological Constant Problem

- ☆ In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- ☆ According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- ☆ Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- ☆ Value of Vacuum Energy Density Makes No Sense:

The Cosmological Constant Problem

- In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- Value of Vacuum Energy Density Makes No Sense: We cannot calculate the vacuum energy density, but the natural particle physics estimate is called the Planck scale — the energy scale at which the effects of quantum gravity are expected to become important. But it is MUCH larger than the observed value:

The Cosmological Constant Problem

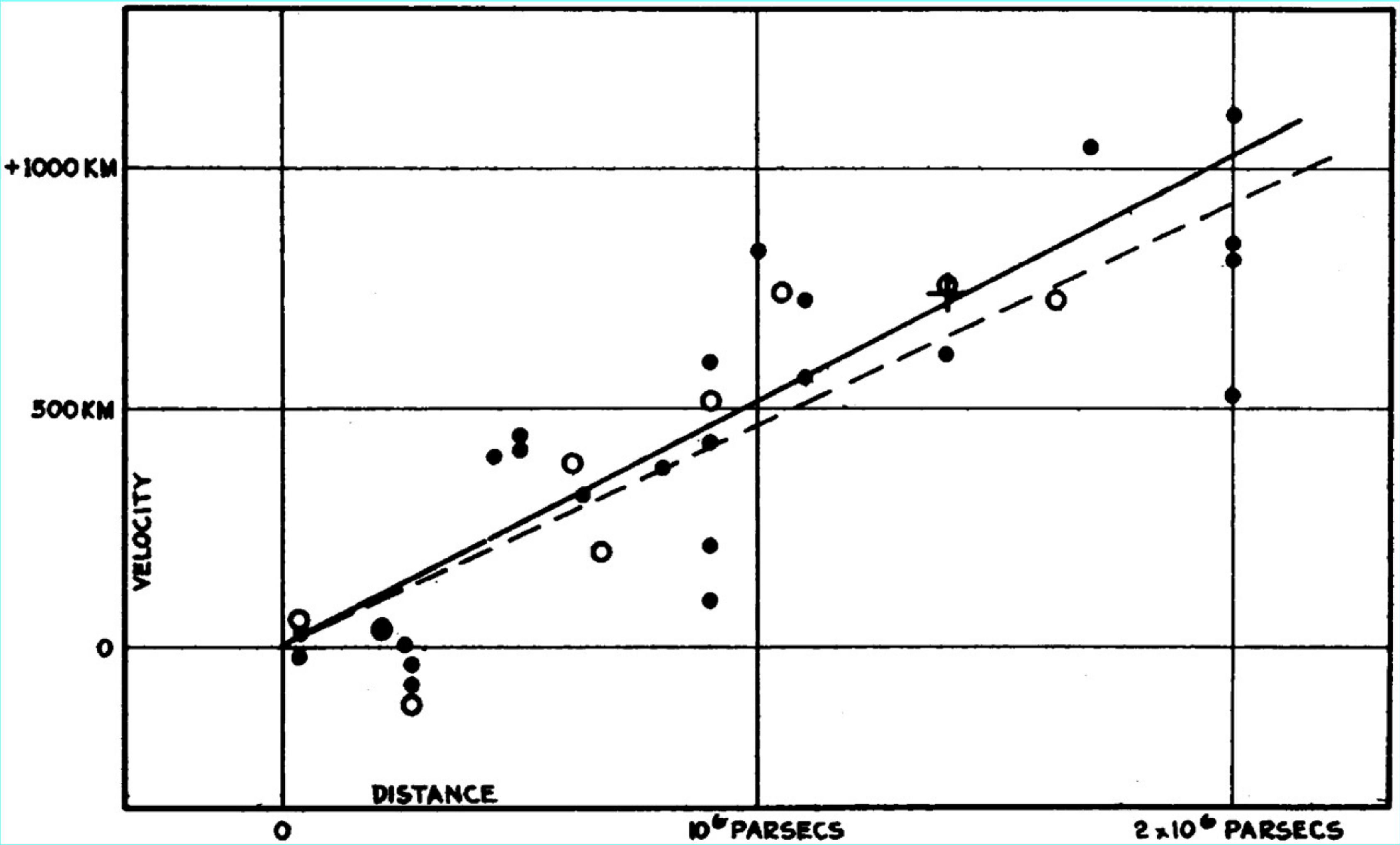
- ☆ In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- ☆ According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- ☆ Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- ☆ Value of Vacuum Energy Density Makes No Sense: We cannot calculate the vacuum energy density, but the natural particle physics estimate is called the Planck scale — the energy scale at which the effects of quantum gravity are expected to become important. But it is MUCH larger than the observed value:

It is larger by 120 orders of magnitude!

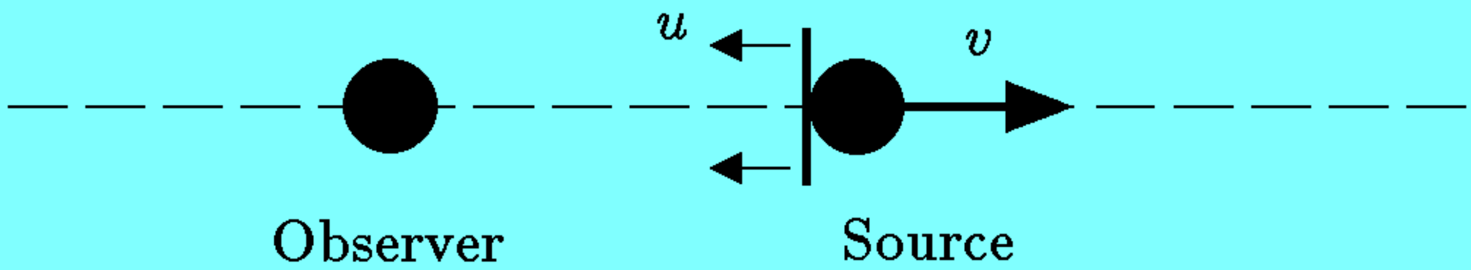
8.286 Lecture 3
September 14, 2022

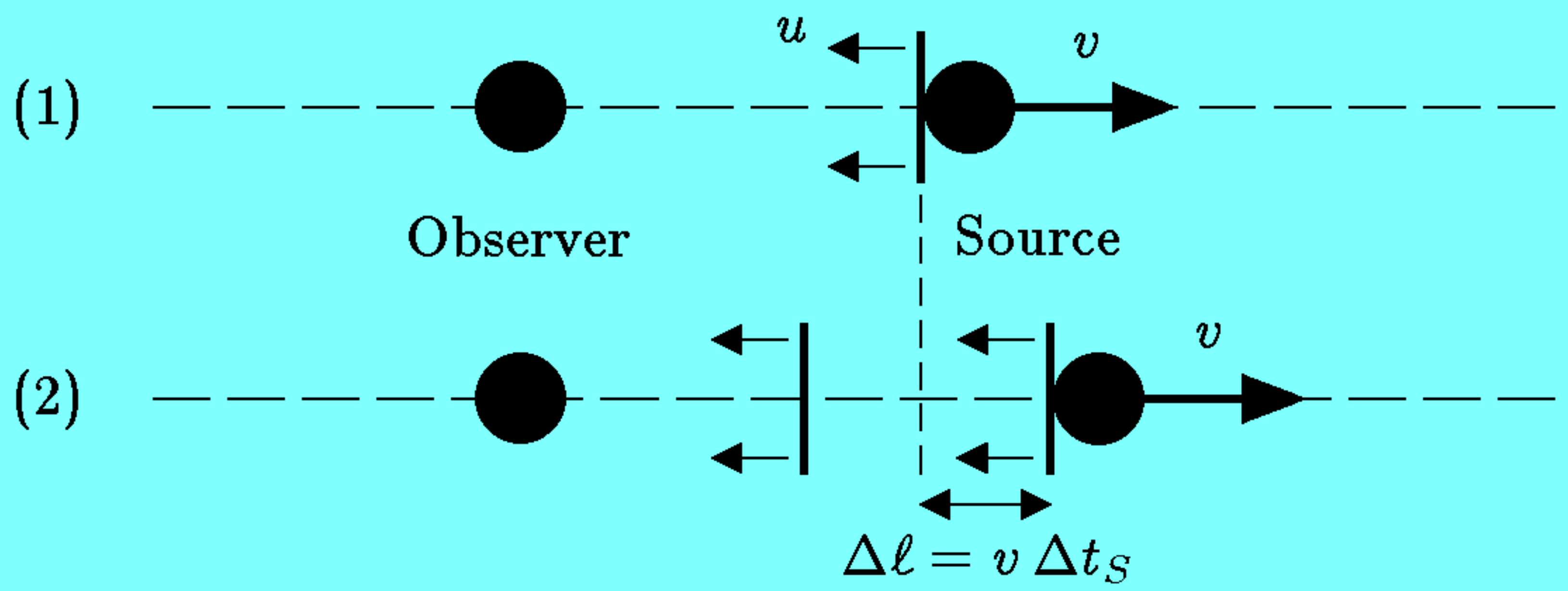
THE DOPPLER EFFECT
and
SPECIAL RELATIVITY

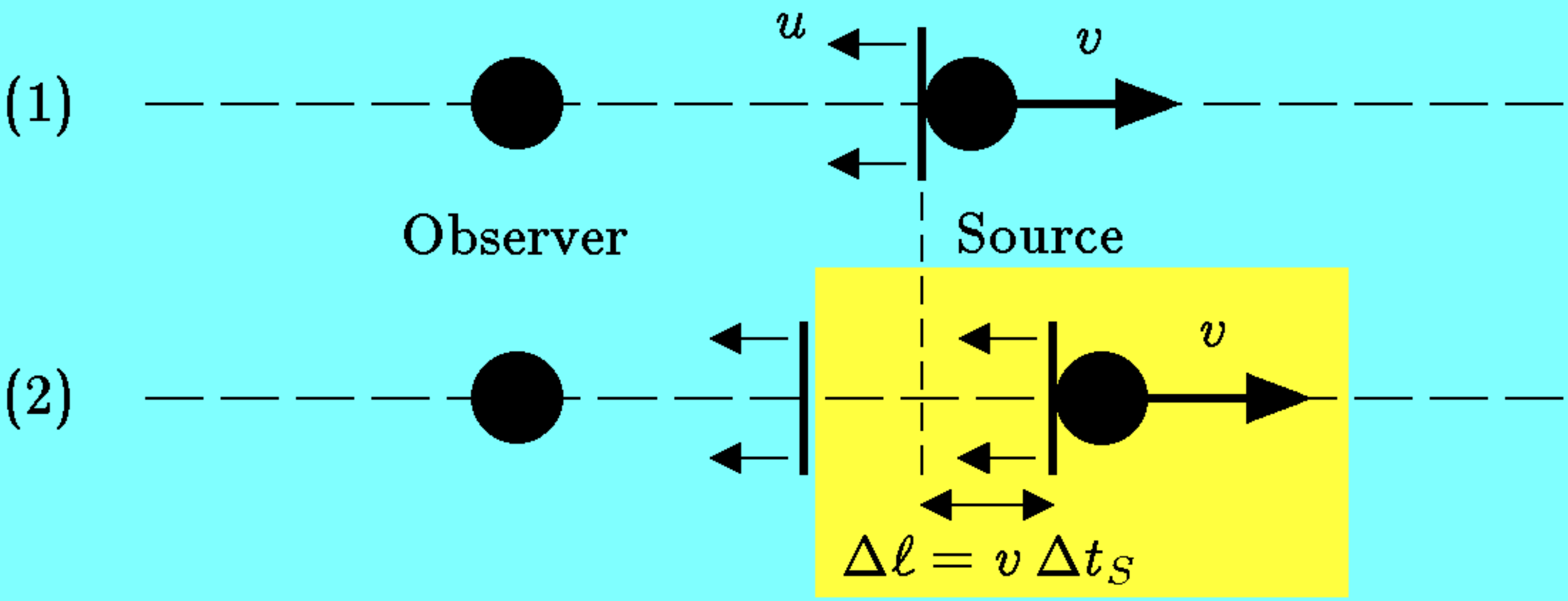
Hubble's Original 1929 Graph



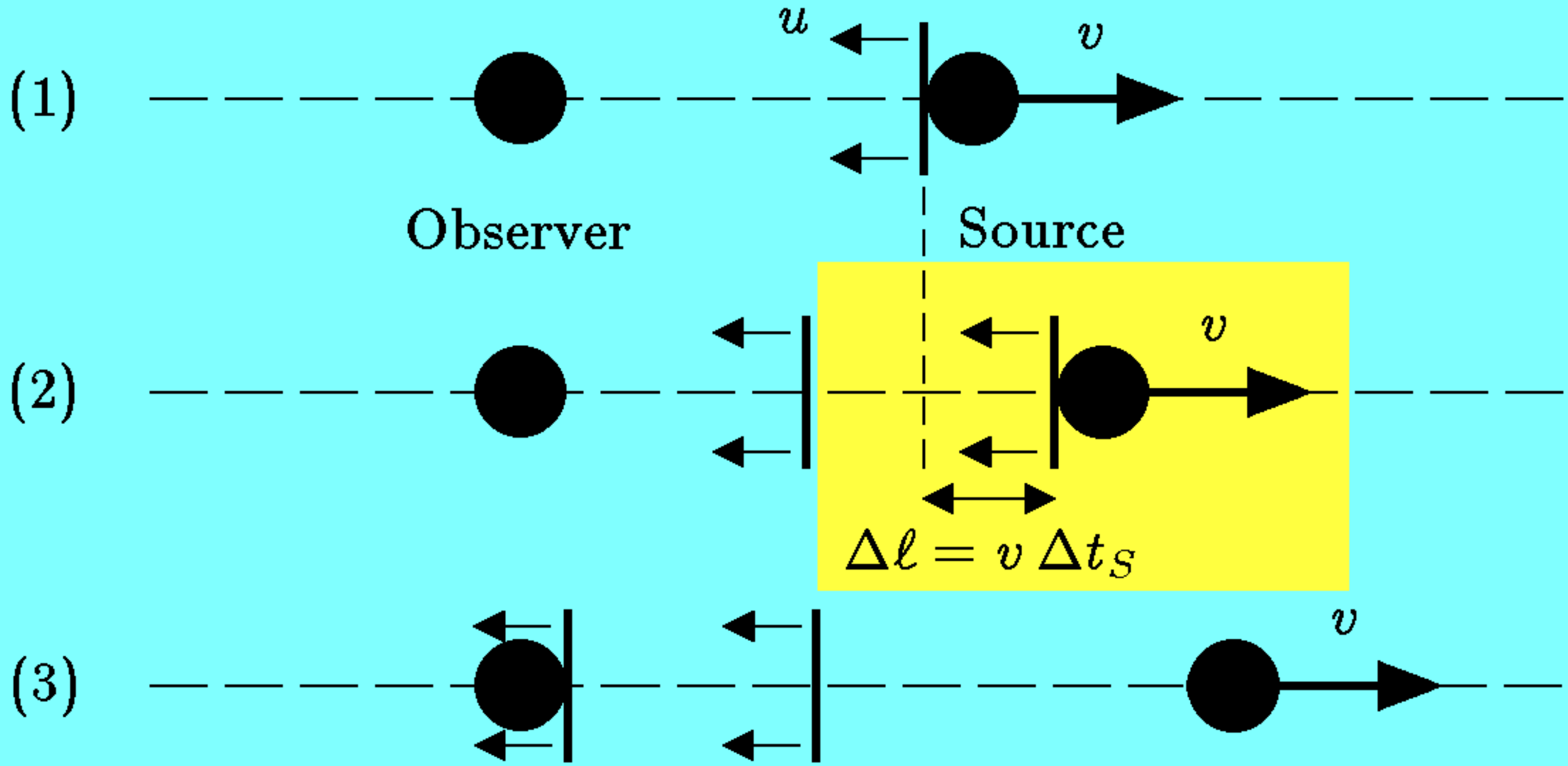
(1)

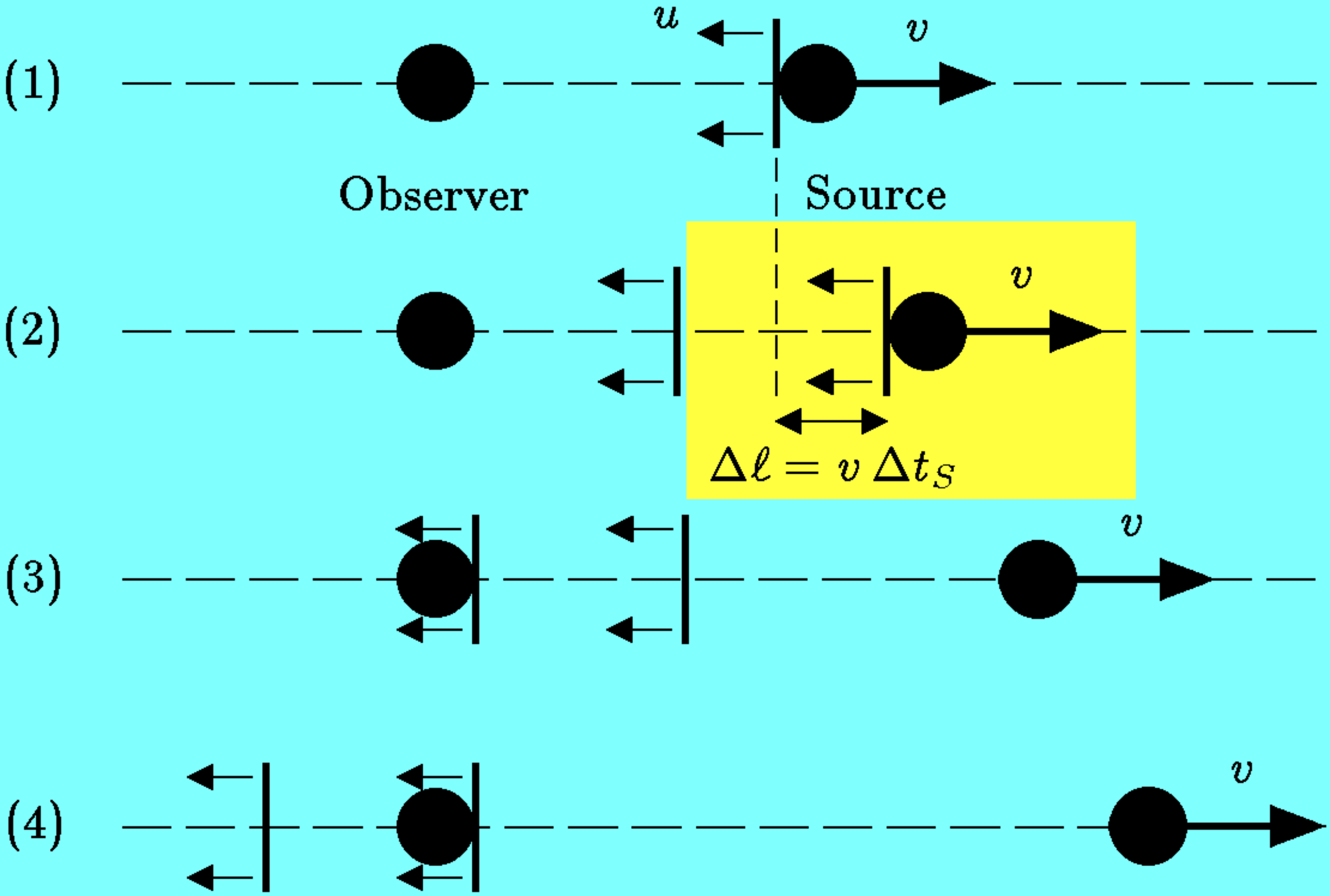



(1)
u
v
Observer
Source
(2)
v
$\Delta\ell = v\,\Delta t_S$

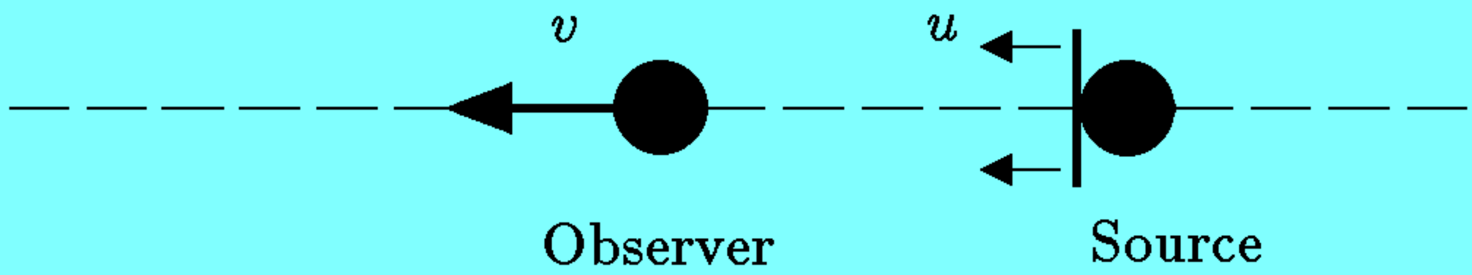


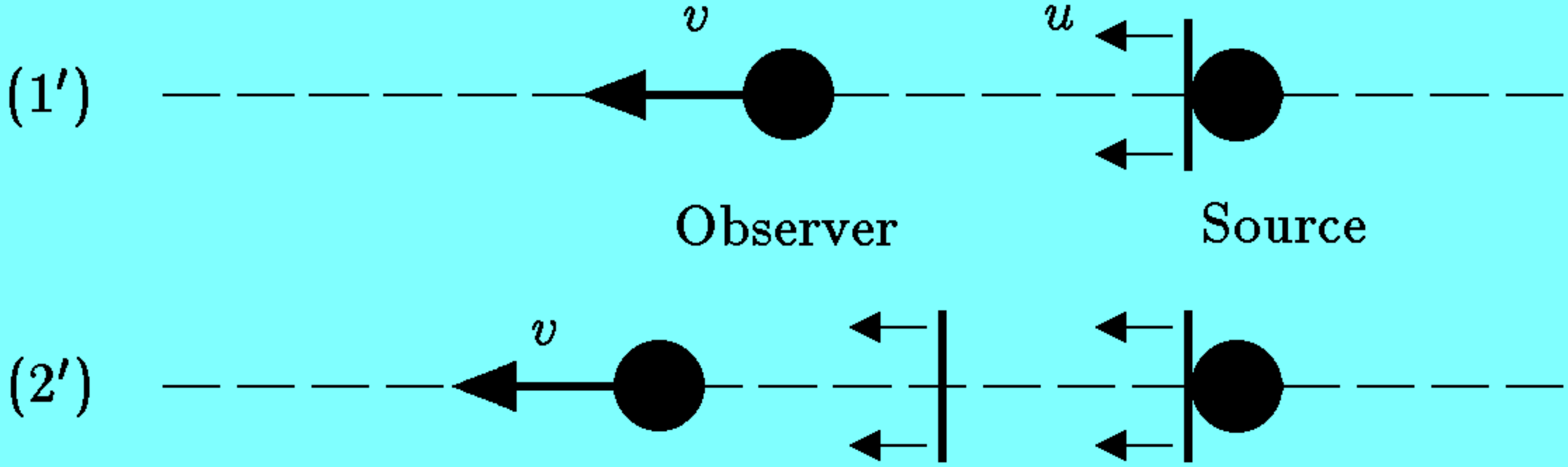
u
v
(1)
Observer
Source
(2)
v
$\Delta\ell = v\,\Delta t_S$

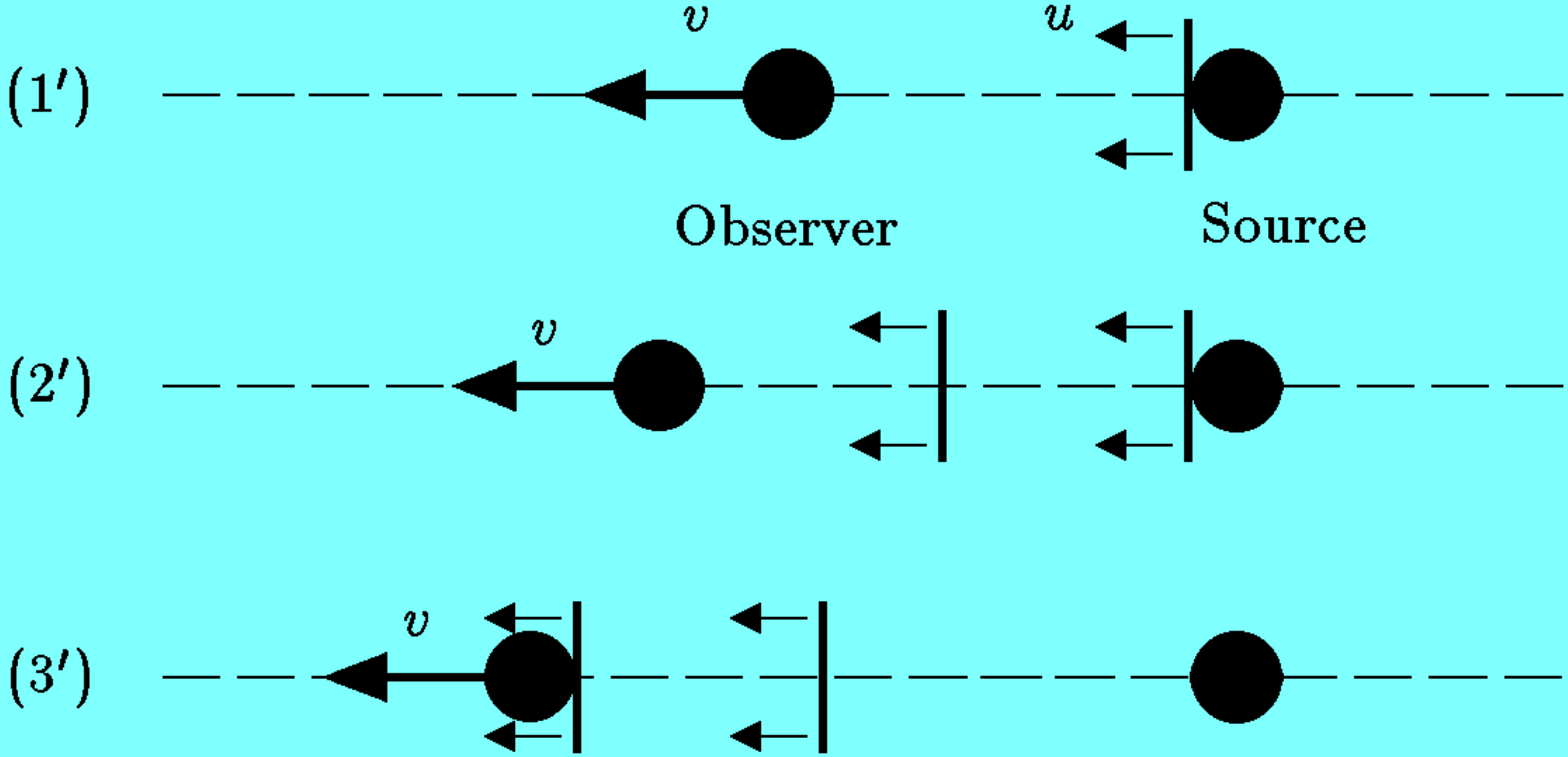

(1)
u
v
Observer
Source
(2)
v
$\Delta \ell = v\,\Delta t_S$
(3)
v

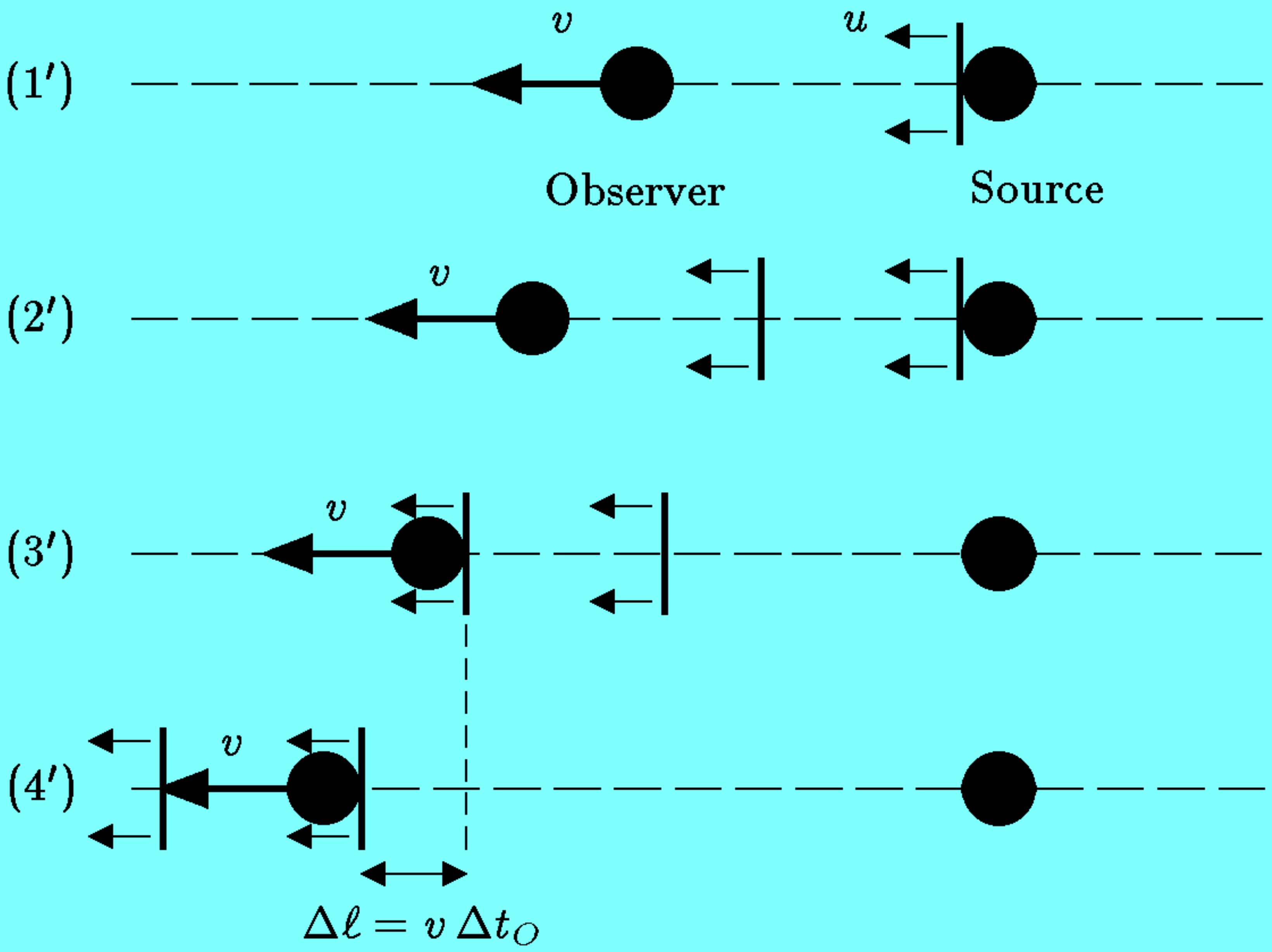

(1)
u
v
Observer
Source
(2)
v
$\Delta\ell = v\,\Delta t_S$
(3)
v
(4)
v

(1′)

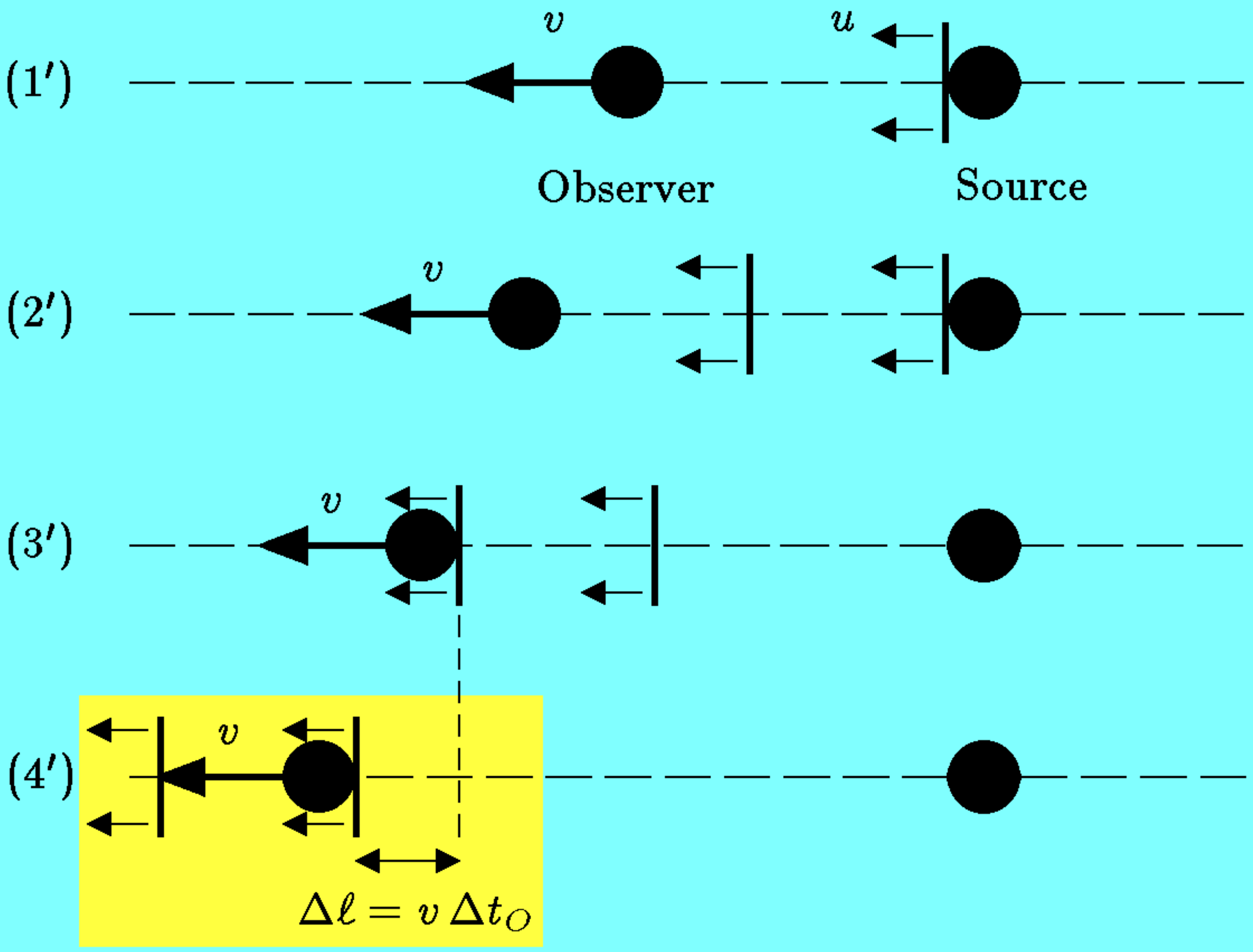



(1′)
v
u
Observer
Source
(2′)
v


(1′)
v
u
Observer
Source
(2′)
v
(3′)
v

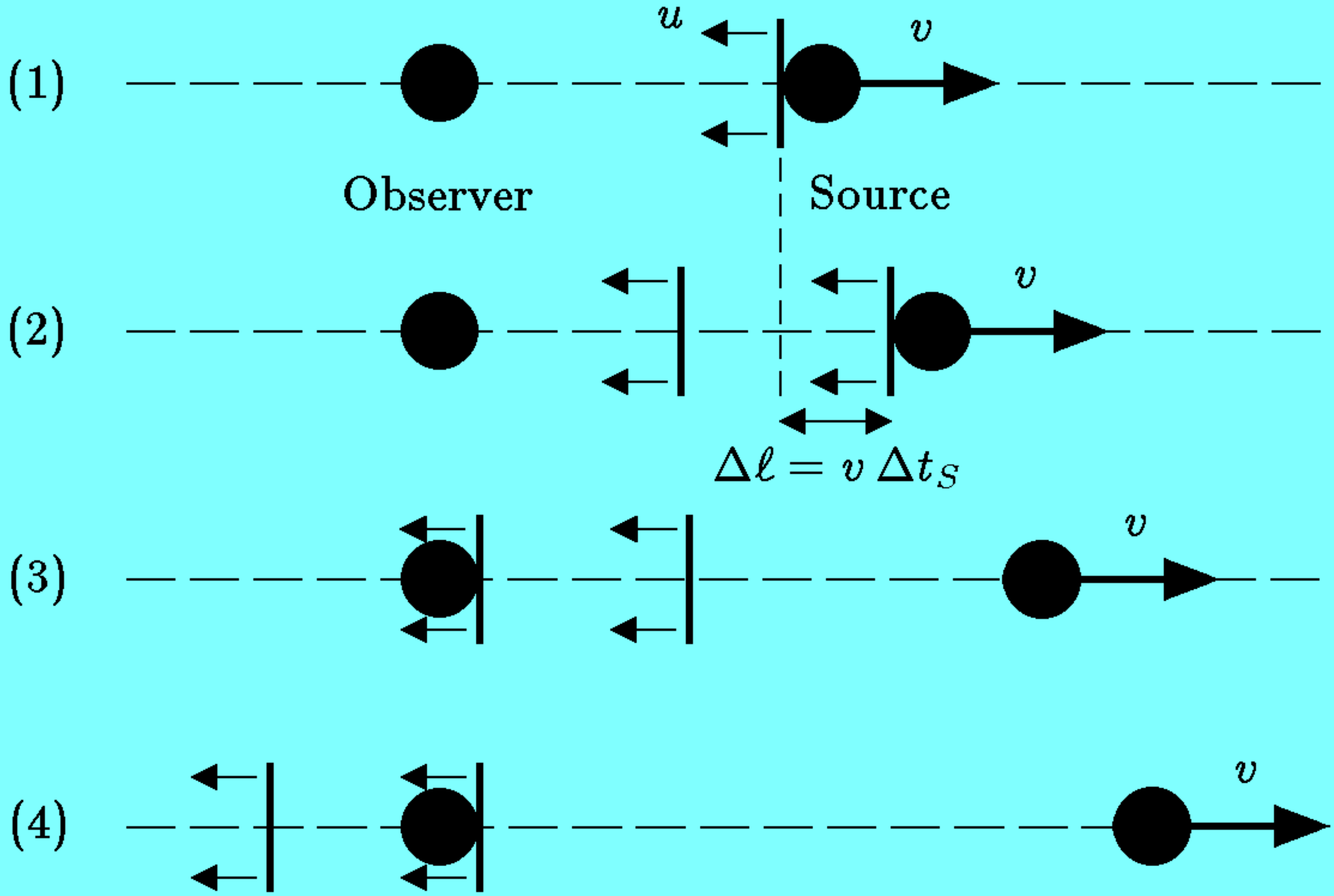


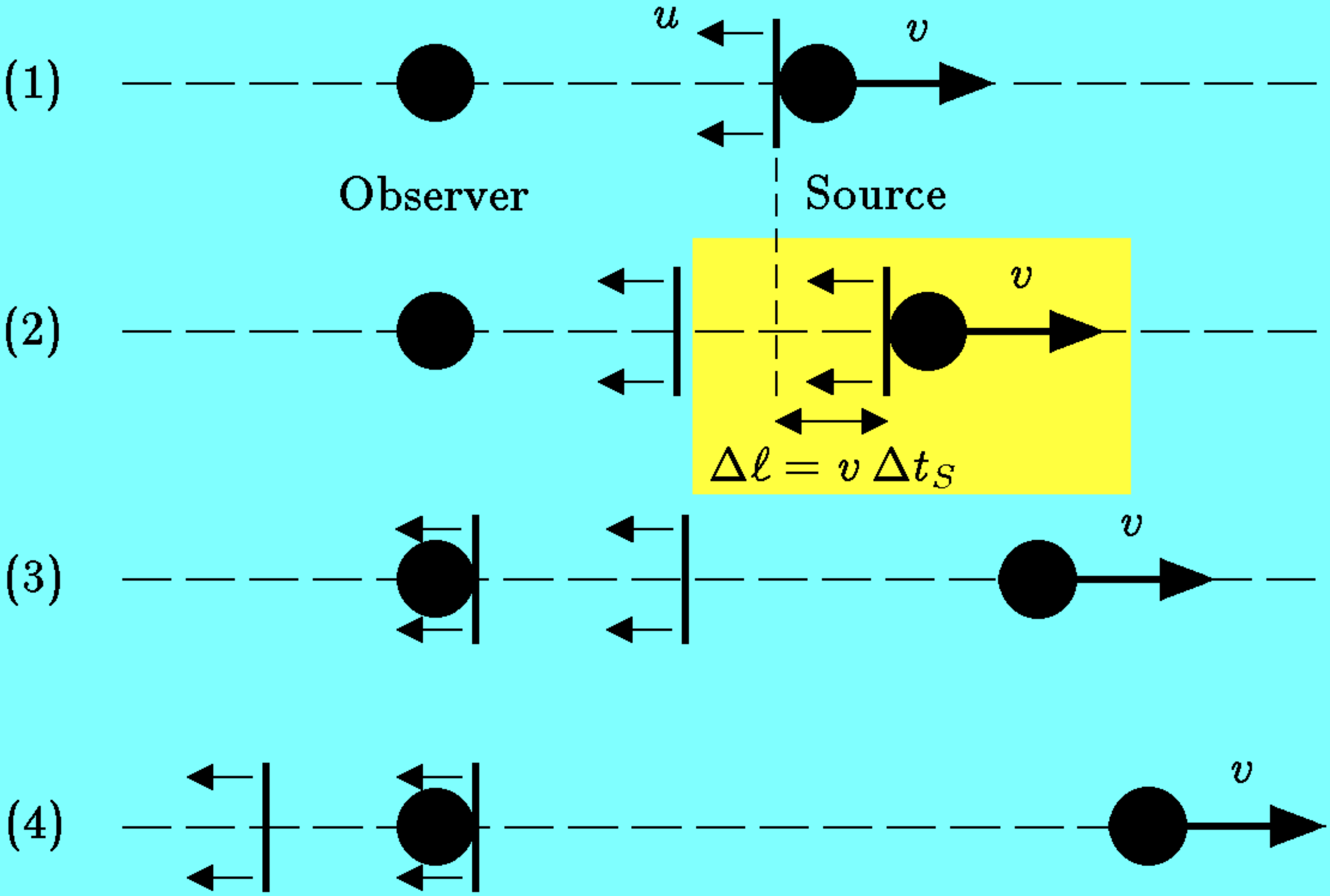
$(1')$
v
u
Observer
Source
$(2')$
v
$(3')$
v
$(4')$
v
$\Delta\ell = v\,\Delta t_O$


$(1')$
v
u
Observer
Source
$(2')$
v
$(3')$
v
$(4')$
v
$\Delta\ell = v\,\Delta t_O$

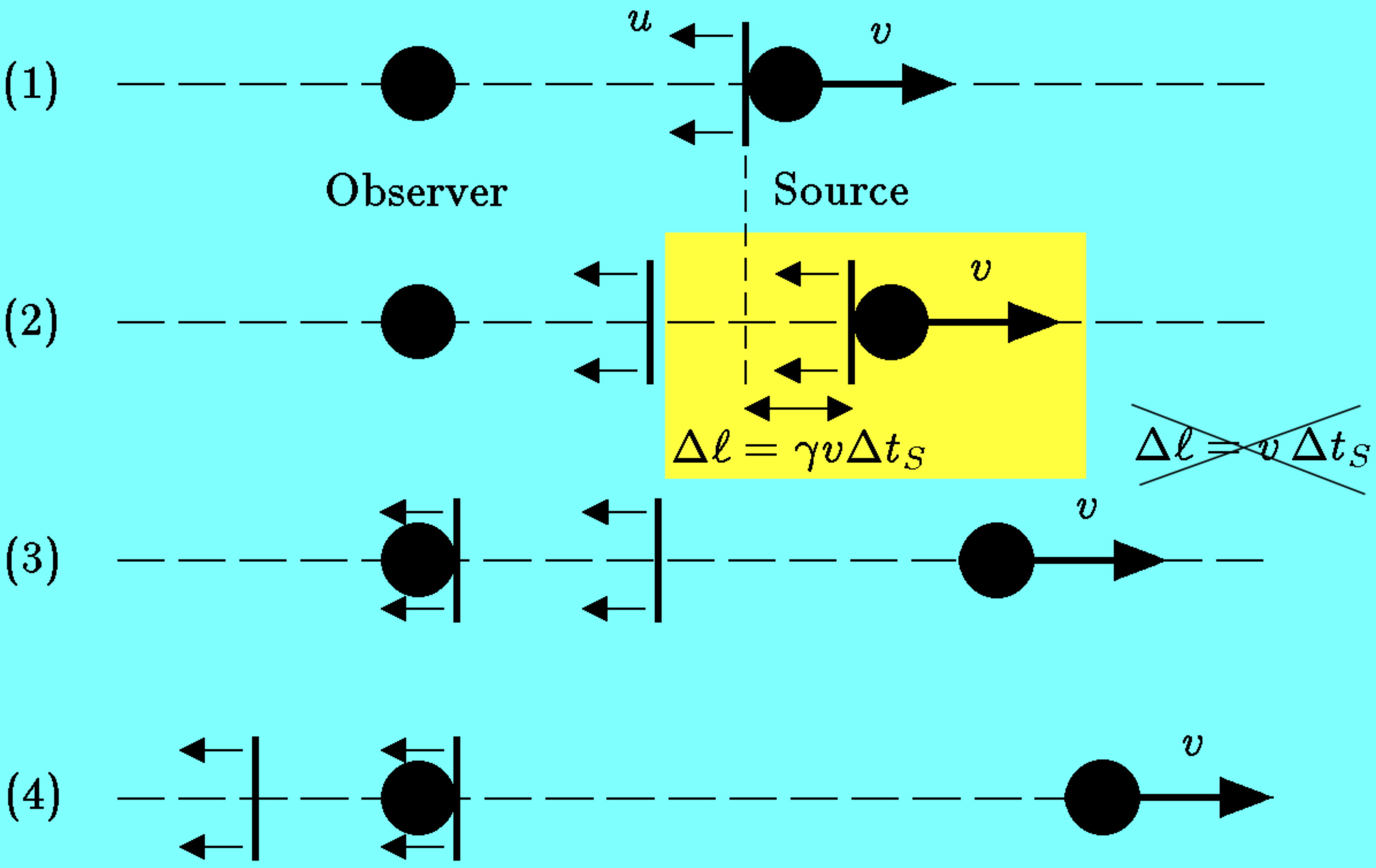
(1) TIME DILATION: Any clock which is moving at speed v relative to a given reference frame will "appear" (to an observer using that reference frame) to run slower than normal by a factor denoted by the Greek letter γ (gamma), and given by

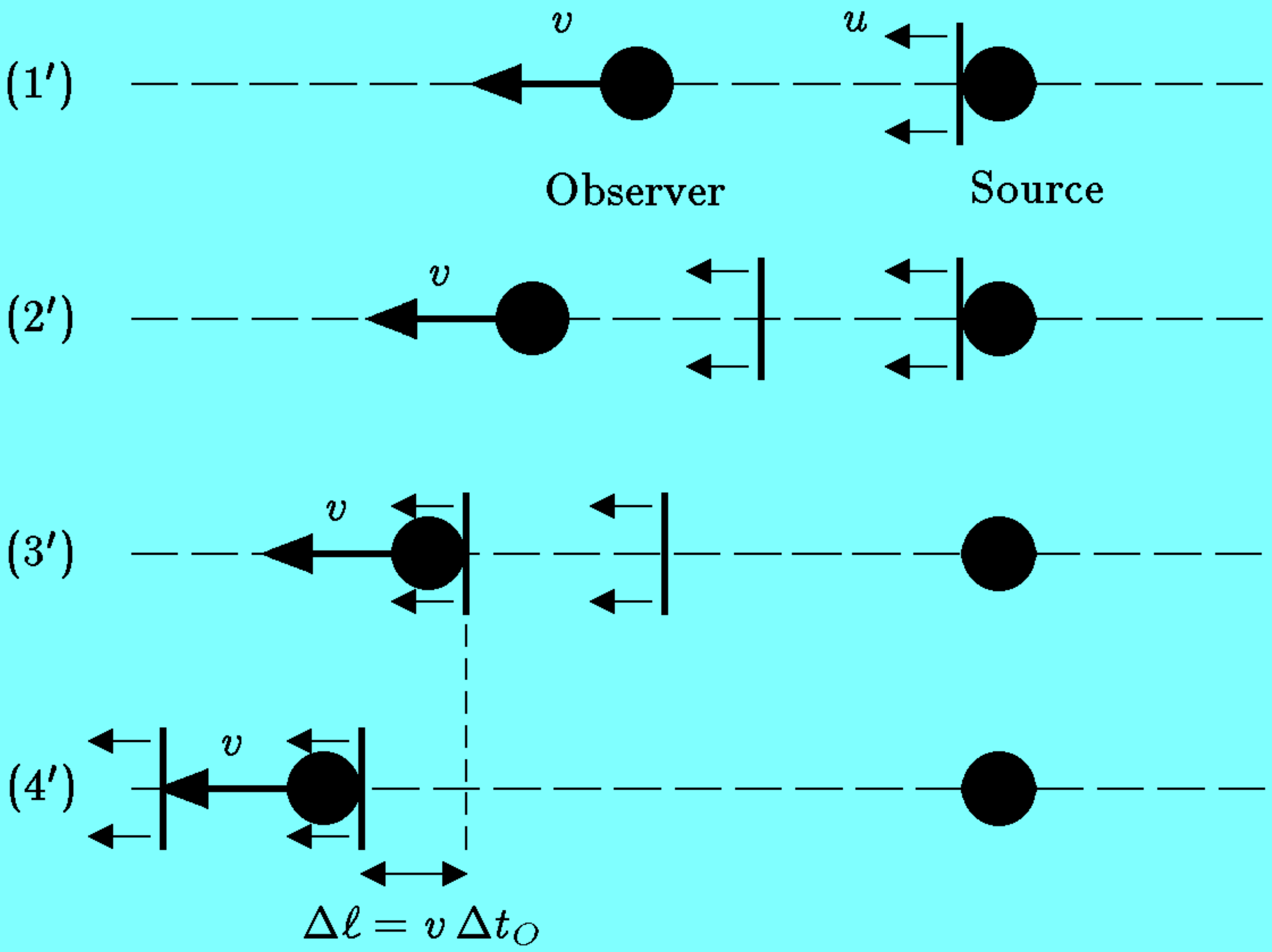
$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}} \; , \qquad \beta \equiv v/c \; . \tag{1.10}$$


(1)
u
v
Observer
Source
(2)
v
$\Delta\ell = v\,\Delta t_S$
(3)
v
(4)
v

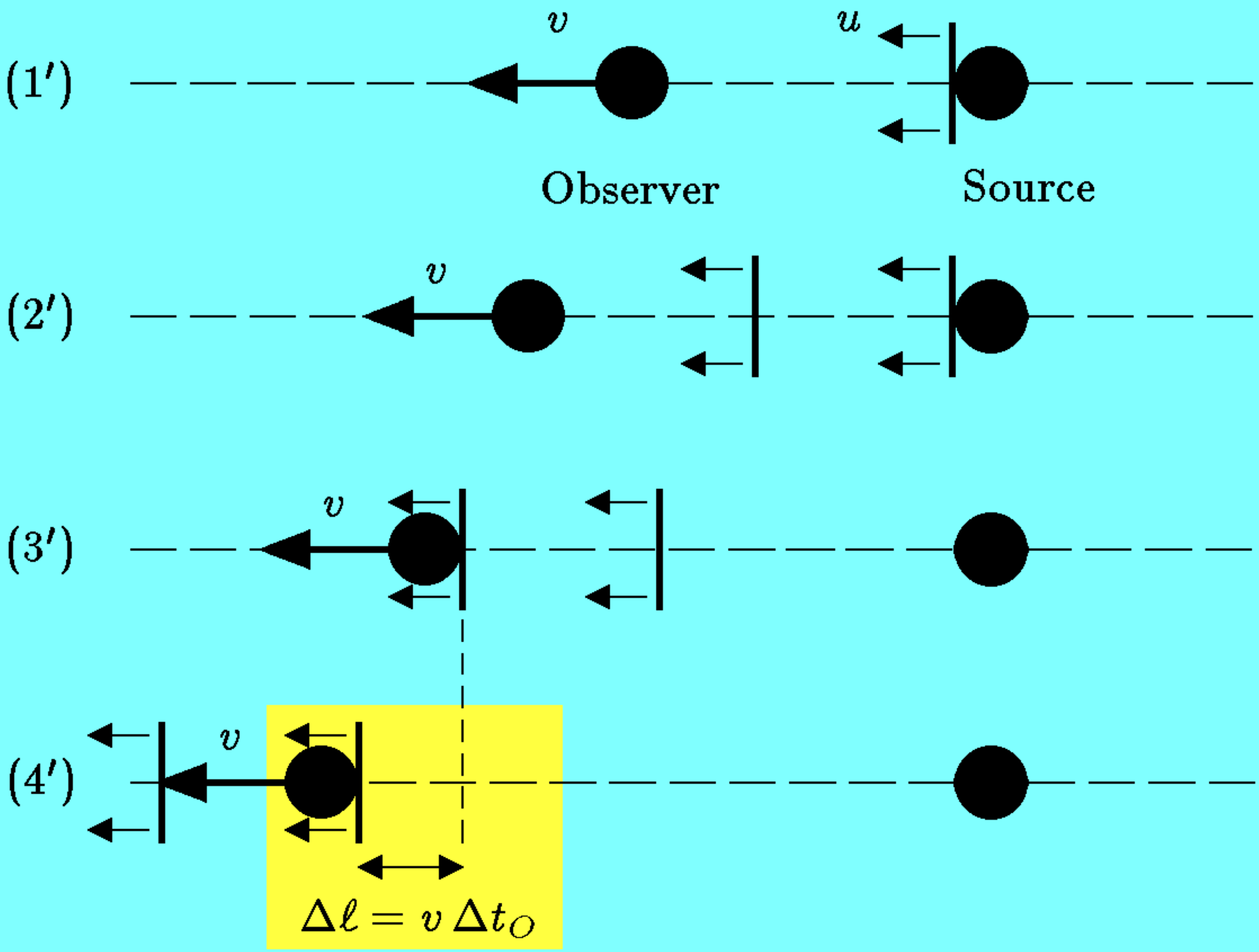


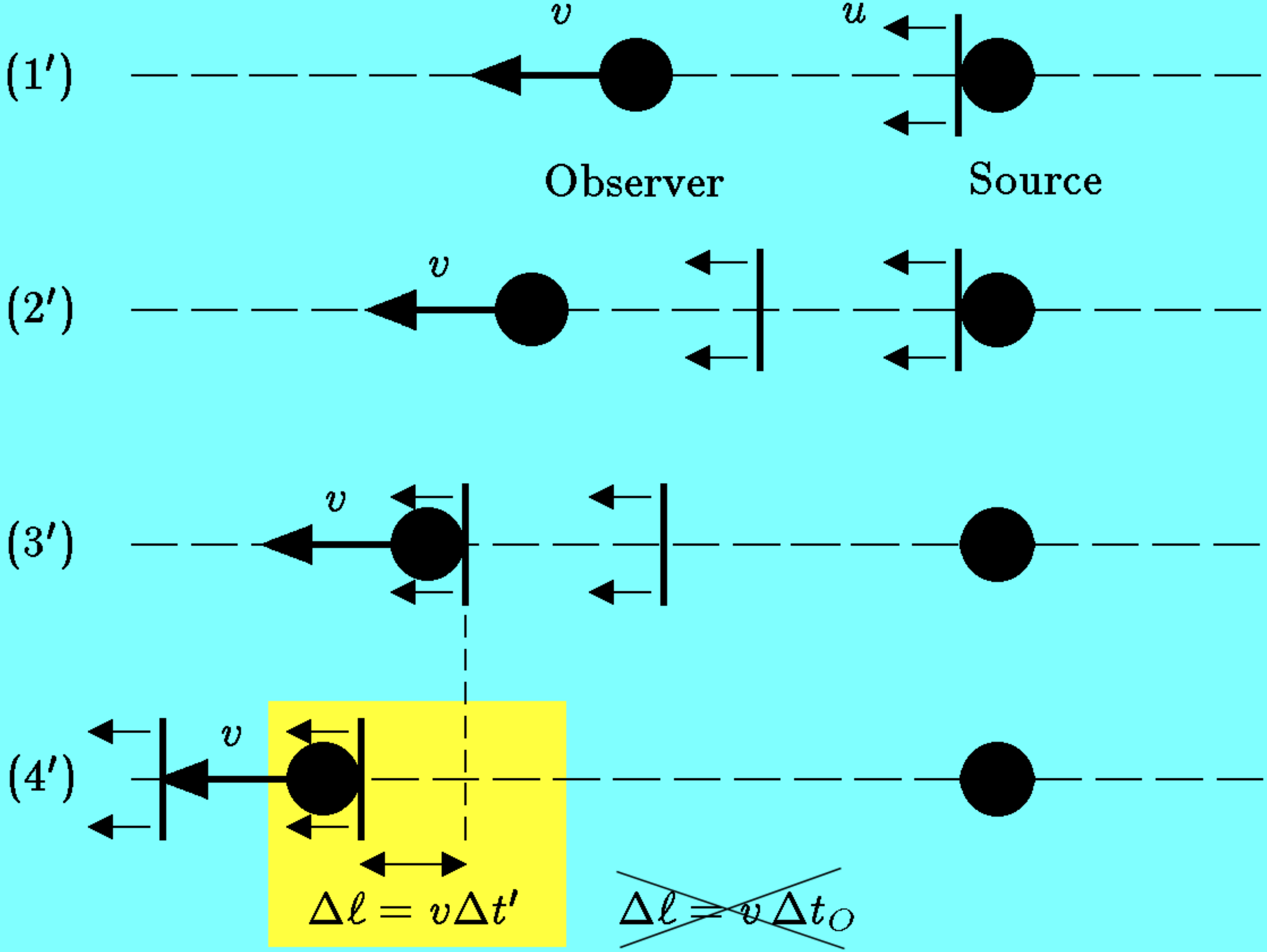
(1)
u
v
Observer
Source
(2)
v
$\Delta\ell = v\,\Delta t_S$
(3)
v
(4)
v


(1)
u
v
Observer
Source
(2)
v
$\Delta\ell = \gamma v \Delta t_S$
$\cancel{\Delta\ell = v\,\Delta t_S}$
(3)
v
(4)
v

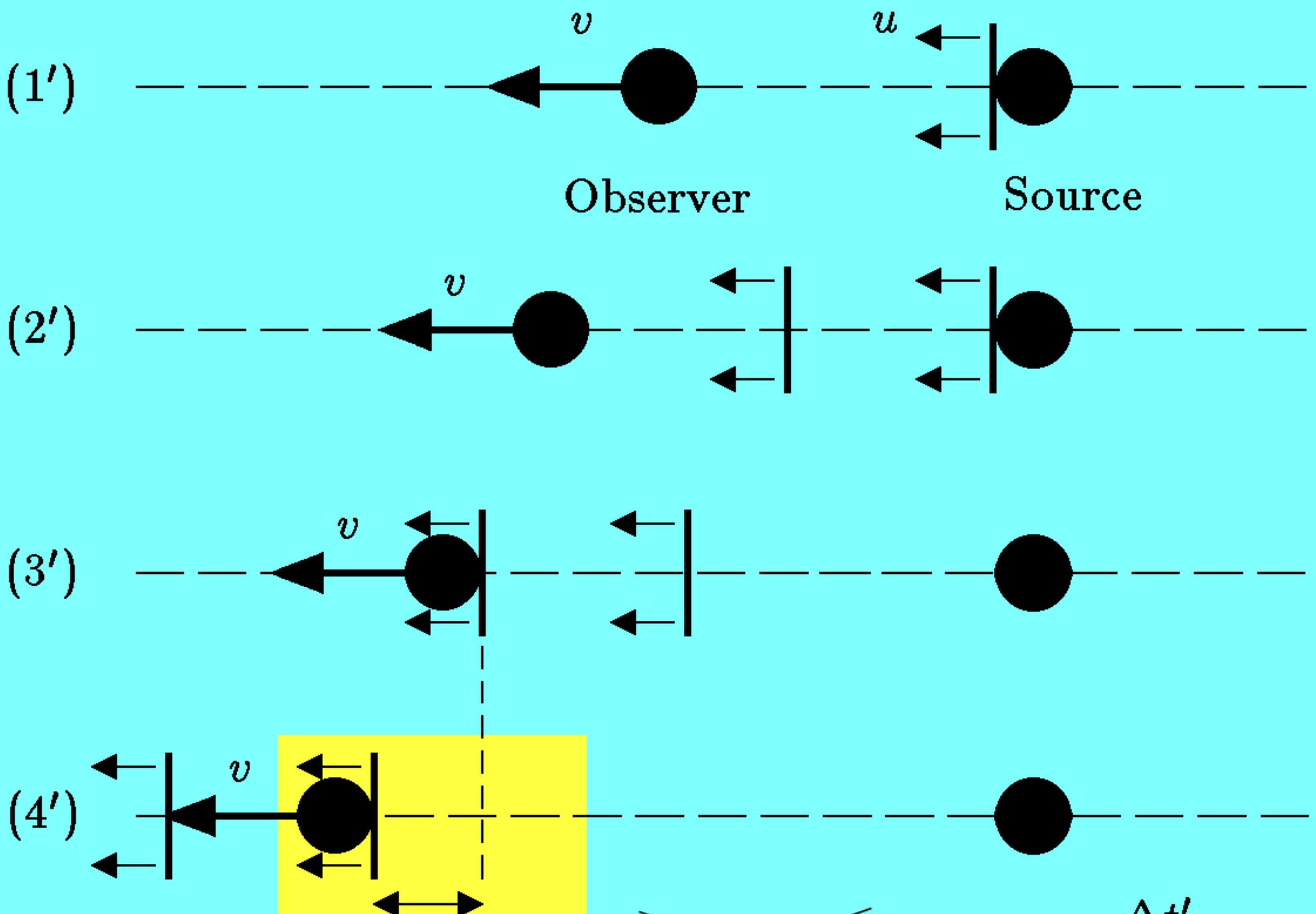


$(1')$
v
u
Observer
Source
$(2')$
v
$(3')$
v
$(4')$
v
$\Delta\ell = v\,\Delta t_O$


(1′)
v
u
Observer
Source
(2′)
v
(3′)
v
(4′)
v
$\Delta\ell = v\,\Delta t_O$



$\Delta t'$ = time between reception of two pulses in frame of the slide.



$\boxed{\Delta\ell = v\Delta t'}$ $\quad \cancel{\Delta\ell = v\Delta t_O} \quad$ $\Delta t_O = \dfrac{\Delta t'}{\gamma}$

$\Delta t'$ = time between reception of two pulses in frame of the slide.

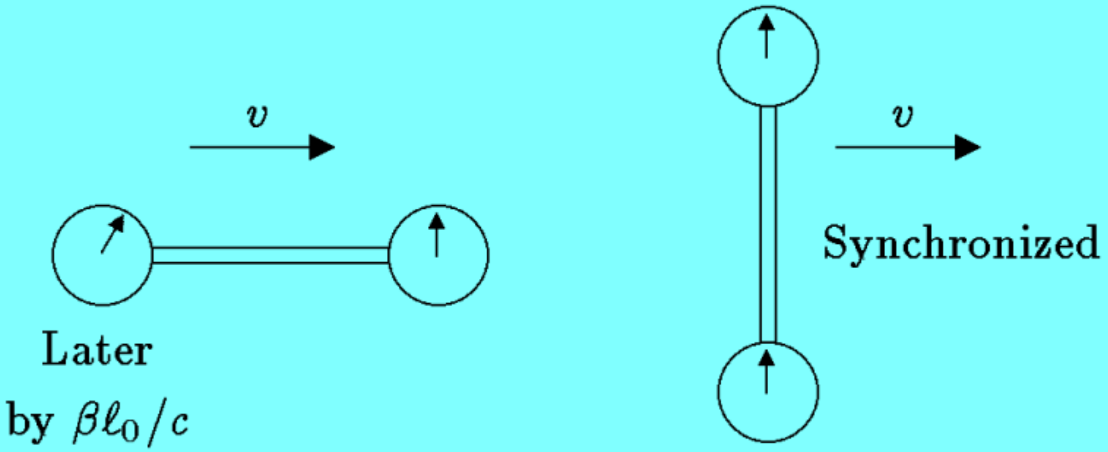
(1) TIME DILATION: Any clock which is moving at speed v relative to a given reference frame will "appear" (to an observer using that reference frame) to run slower than normal by a factor denoted by the Greek letter γ (gamma), and given by

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}} \; , \qquad \beta \equiv v/c \; . \tag{1.10}$$

(2) LORENTZ-FITZGERALD CONTRACTION: Any rod which is moving at a speed v along its length relative to a given reference frame will "appear" (to an observer using that reference frame) to be shorter than its normal length by the same factor γ. A rod which is moving perpendicular to its length does not undergo a change in apparent length.

Contracted by γ Uncontracted

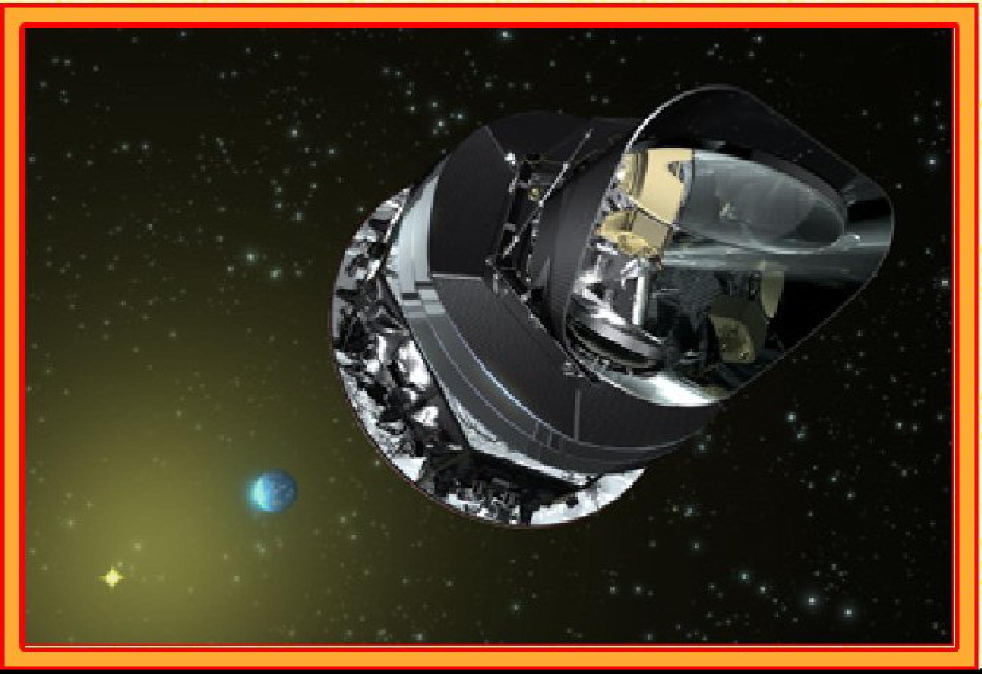
(3) RELATIVITY OF SIMULTANEITY: Suppose a rod which has rest length ℓ_0 is equipped with a clock at each end. The clocks can be synchronized in the rest frame of the system by using light pulses. (That is, a light pulse can be sent out from the center, and the clocks at both ends can be started when they receive the pulses.) If the system moves at speed v along its length, then the trailing clock will "appear" to read a time which is later than the leading clock by an amount $\beta\ell_0/c$. If, on the other hand, the system moves perpendicular to its length, then the synchronization of the clocks is not disturbed.



INFLATIONARY COSMOLOGY:

IS OUR UNIVERSE PART OF A MULTIVERSE?

PART 3



— Alan Guth —

Massachusetts Institute of Technology

8.286 Lecture 4
September 19, 2022

Inflation Suggests a Multiverse

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Roughly speaking, inflation is driven by a metastable state, which decays with some half-life.

After one half-life, half of the inflating material has become normal, noninflating matter, but the half that remains has continued to expand exponentially. It is vastly larger than it was at the beginning.

Once started, the inflation goes on FOREVER, with pieces of the inflating region breaking off and producing "pocket universes."

Inflation Suggests a Multiverse

Almost all detailed models of inflation lead to "eternal inflation," and hence to a multiverse.

Roughly speaking, inflation is driven by a metastable state, which decays with some half-life.

After one half-life, half of the inflating material has become normal, noninflating matter, but the half that remains has continued to expand exponentially. It is vastly larger than it was at the beginning.

Once started, the inflation goes on FOREVER, with pieces of the inflating region breaking off and producing "pocket universes."

We would be living in one of the infinity of pocket universes.

The Cosmological Constant Problem

- ☆ In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- ☆ According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- ☆ Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.

The Cosmological Constant Problem

- In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- Value of Vacuum Energy Density Makes No Sense:

The Cosmological Constant Problem

- ☆ In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- ☆ According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- ☆ Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- ☆ Value of Vacuum Energy Density Makes No Sense: We cannot calculate the vacuum energy density, but the natural particle physics estimate is called the Planck scale — the energy scale at which the effects of quantum gravity are expected to become important. But it is MUCH larger than the observed value:

Review from previous lecture

The Cosmological Constant Problem

- ☆ In 1998, two groups of astronomers discovered that for the past 5–6 billion years, the expansion of the universe has been accelerating.
- ☆ According to GR, this requires a repulsive gravity material (i.e., a negative pressure material), which is dubbed "Dark Energy".
- ☆ Simplest explanation: dark energy is vacuum energy — the energy density of empty space. The physicist's vacuum is far from empty, so a nonzero energy density is expected.
- ☆ Value of Vacuum Energy Density Makes No Sense: We cannot calculate the vacuum energy density, but the natural particle physics estimate is called the Planck scale — the energy scale at which the effects of quantum gravity are expected to become important. But it is MUCH larger than the observed value:

It is larger by 120 orders of magnitude!

The Multiverse and the Cosmological Constant Problem

- ☆ The multiverse offers a possible (although controversial) explanation of why the energy density of the vacuum (equivalent to the cosmological constant) is 120 orders of magnitude smaller than the (expected) Planck scale.

The Multiverse and the Cosmological Constant Problem

- The multiverse offers a possible (although controversial) explanation of why the energy density of the vacuum (equivalent to the cosmological constant) is 120 orders of magnitude smaller than the (expected) Planck scale.
- If there are 10^{500} or more different types of vacuum (as in string theory), there will be many with energy densities in the range we observe, although they will be only a tiny fraction of types of vacuum.

The Multiverse and the Cosmological Constant Problem

- ☆ The multiverse offers a possible (although controversial) explanation of why the energy density of the vacuum (equivalent to the cosmological constant) is 120 orders of magnitude smaller than the (expected) Planck scale.
- ☆ If there are 10^{500} or more different types of vacuum (as in string theory), there will be many with energy densities in the range we observe, although they will be only a tiny fraction of types of vacuum.
- ☆ But why should we find ourselves in such an extremely rare kind of vacuum?

The Multiverse and the Cosmological Constant Problem

- The multiverse offers a possible (although controversial) explanation of why the energy density of the vacuum (equivalent to the cosmological constant) is 120 orders of magnitude smaller than the (expected) Planck scale.
- If there are 10^{500} or more different types of vacuum (as in string theory), there will be many with energy densities in the range we observe, although they will be only a tiny fraction of types of vacuum.
- But why should we find ourselves in such an extremely rare kind of vacuum?
- The vacuum energy affects cosmic evolution: if it is too large and positive, the universe flies apart too fast for galaxies to form. If too large and negative, the universe implodes.

The Multiverse and the Cosmological Constant Problem

- The multiverse offers a possible (although controversial) explanation of why the energy density of the vacuum (equivalent to the cosmological constant) is 120 orders of magnitude smaller than the (expected) Planck scale.
- If there are 10^{500} or more different types of vacuum (as in string theory), there will be many with energy densities in the range we observe, although they will be only a tiny fraction of types of vacuum.
- But why should we find ourselves in such an extremely rare kind of vacuum?
- The vacuum energy affects cosmic evolution: if it is too large and positive, the universe flies apart too fast for galaxies to form. If too large and negative, the universe implodes.
- It is therefore plausible that life only forms in those pocket universes with incredibly small vacuum energies, so all living beings would observe a small vacuum energy. (Anthropic principle, or observational selection effect.)

SUMMARY

The Inflationary Paradigm is in Great Shape!

- ☆ Explains large scale uniformity.
- ☆ Predicts the mass density of the universe to better than 1% accuracy.
- ☆ Explains the ripples we see in the cosmic background radiation as the result of quantum fluctuations.

Three Strong Winds Blowing Us Towards the Multiverse — a diverse multiverse where selection effects play an important role

1) Theoretical Cosmology: Almost all inflationary models are eternal into the future. Once inflation starts, it never stops, but goes on forever producing pocket universes.

2) Observational Astronomy: Astronomers have discovered that the universe is accelerating, which probably indicates a vacuum energy that is nonzero, but incredibly much smaller than we can understand. Why should this happen?

3) String Theory: String theorists mostly agree that string theory has no unique vacuum, but instead a landscape of perhaps 10^{500} or more long-lived metastable states, any of which could serve as the substrate for a pocket universe, including our own. This situation allows an "anthropic" argument: perhaps we see an incredibly small vacuum energy density because conscious beings only form in those parts of the multiverse where the vacuum energy density is incredibly small.

8.286 Lecture 4
September 19, 2022

THE KINEMATICS of a HOMOGENEOUSLY EXPANDING UNIVERSE

Hubble's Law

$$\boxed{v = Hr\ .}$$

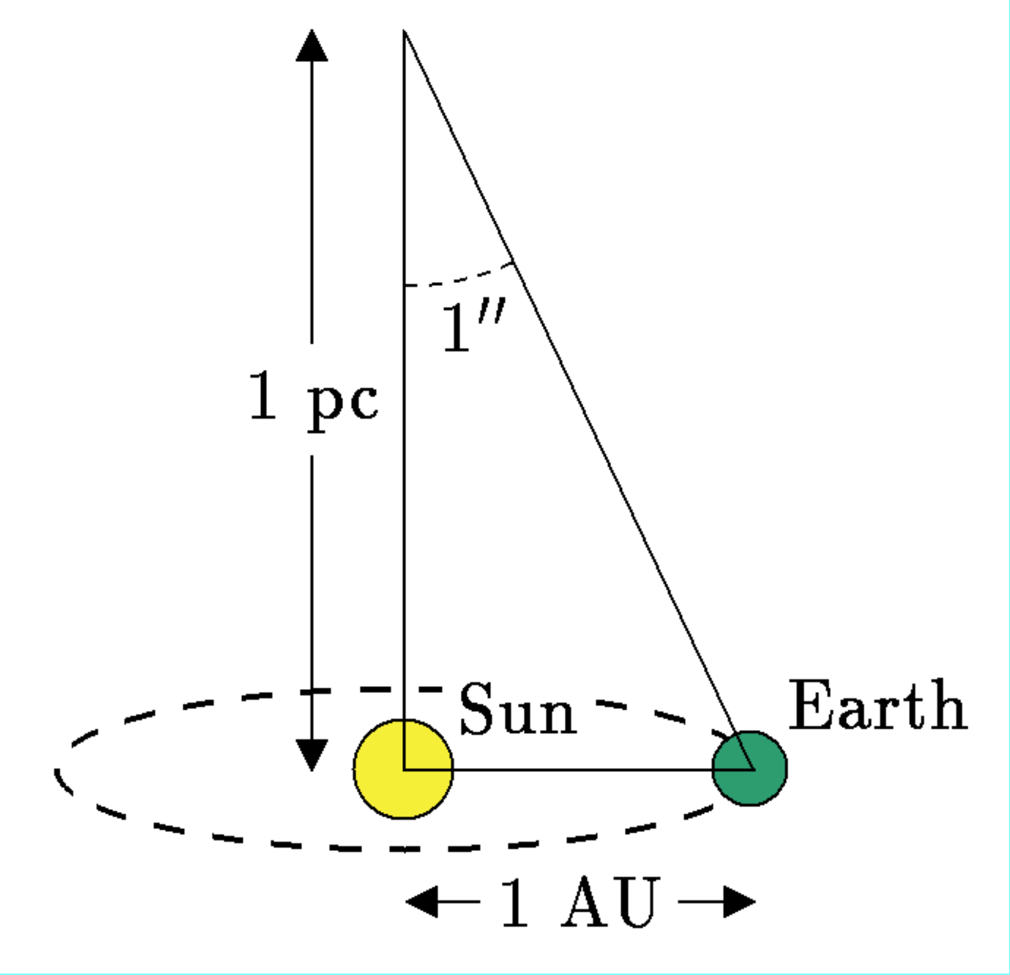
Here

$$v \equiv \text{recession velocity}\ ,$$
$$H \equiv \text{Hubble expansion rate}\ ,$$

and

$$r \equiv \text{distance to galaxy}\ .$$

The Parsec



This slide was not shown, but the work was done on the blackboard.

Units for the Hubble Expansion Rate

$$v = Hr \quad \Longrightarrow \quad [H] = [v]/[r] = (L/T)/L = 1/T .$$

Astronomers invariably think in terms of velocity/distance, which they measure in km-s^{-1}-Mpc^{-1}.

1 pc = 3.2616 light-yr; Mpc = megaparsec = 10^6 pc.

Relation to inverse time:

$$\frac{1}{10^{10}\ \text{yr}} = 97.8\ \text{km-s}^{-1}\text{-Mpc}^{-1} .$$

Homogeneity and Hubble's Law

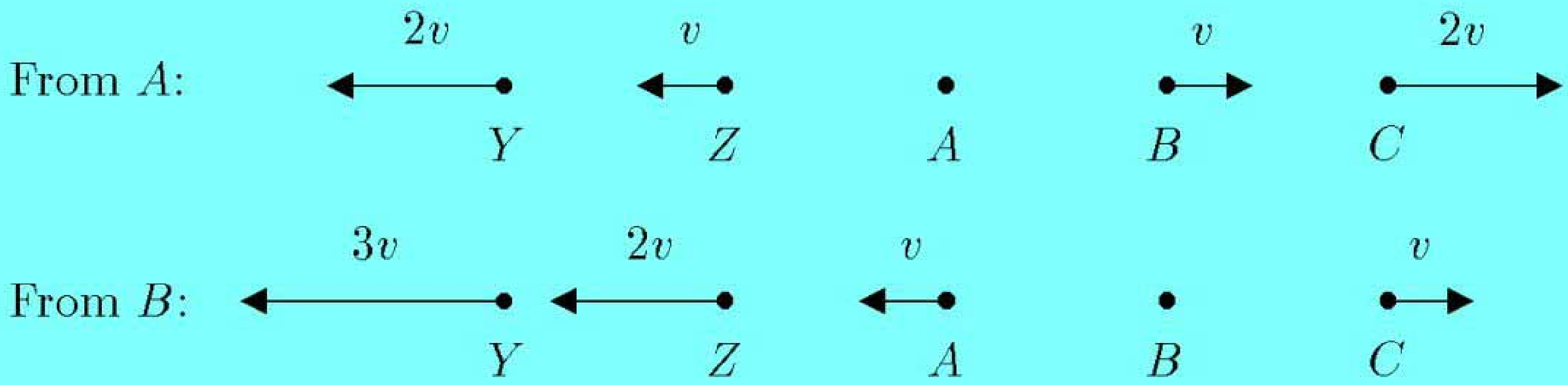
☆ Does Hubble's law imply that we are in the center of the universe?

Homogeneity and Hubble's Law

☆ Does Hubble's law imply that we are in the center of the universe? No.

Homogeneity and Hubble's Law

- ☆ Does Hubble's law imply that we are in the center of the universe? No.
- ☆ As Weinberg explains it in *The First Three Minutes:*



This slide was not shown, but the work was done on the blackboard.

Comoving Coordinates

☆ If the Earth kept getting larger, uniformly, would be have to keep redrawing the map?

This slide was not shown, but the work was done on the blackboard.

Comoving Coordinates

- If the Earth kept getting larger, uniformly, would be have to keep redrawing the map?
- No. Any map has a scale marked in the corner someplace: e.g., 1 inch = 1,000 miles. If the Earth kept getting larger, uniformly, we could keep the map and continuously change the scale.

This slide was not shown, but the work was done on the blackboard.

Comoving Coordinates

- If the Earth kept getting larger, uniformly, would be have to keep redrawing the map?
- No. Any map has a scale marked in the corner someplace: e.g., 1 inch = 1,000 miles. If the Earth kept getting larger, uniformly, we could keep the map and continuously change the scale. That is what we do in cosmology.

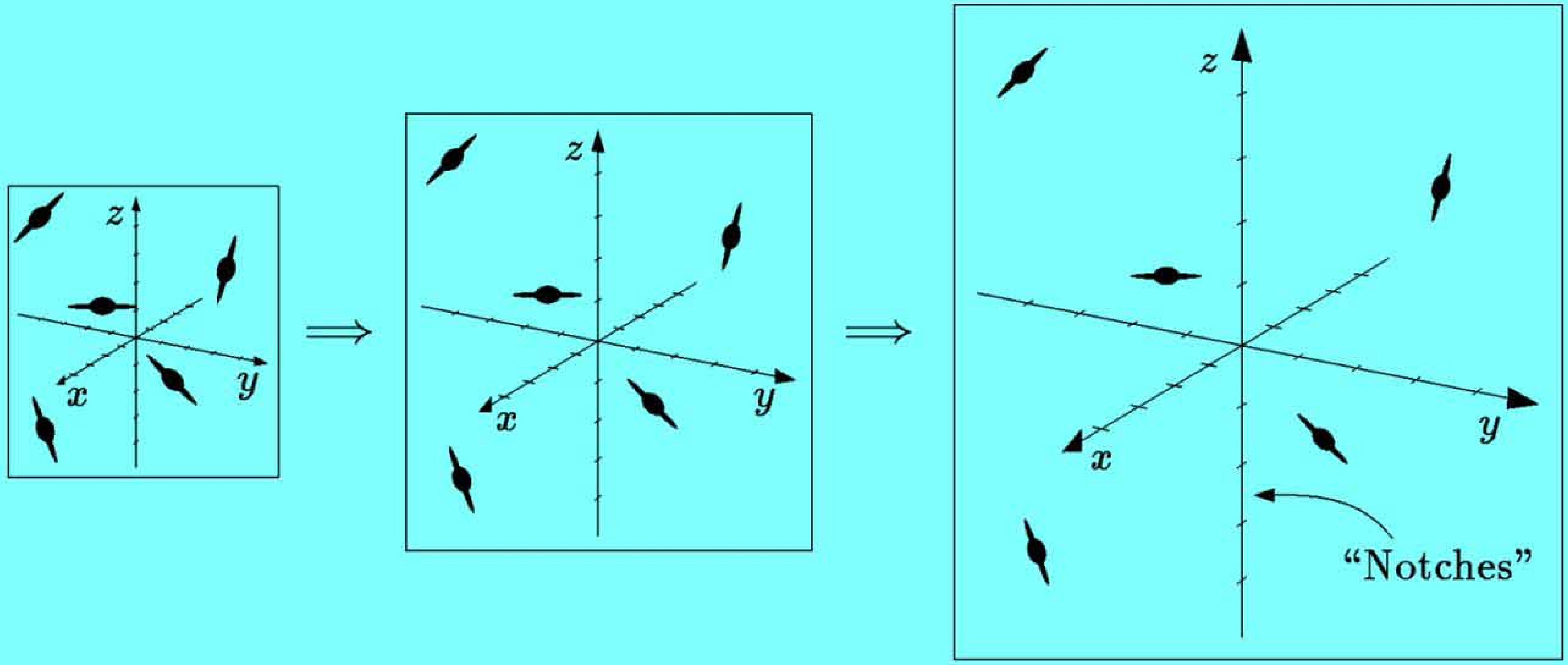
This slide was not shown, but the work was done on the blackboard.

Comoving Coordinates

- If the Earth kept getting larger, uniformly, would be have to keep redrawing the map?
- No. Any map has a scale marked in the corner someplace: e.g., 1 inch = 1,000 miles. If the Earth kept getting larger, uniformly, we could keep the map and continuously change the scale. That is what we do in cosmology.
- We imagine a fixed 3D map of the universe, with distances marked in some arbitrary unit: I call them "notches," to make it clear that they have no fixed meaning in terms of any standard units of length. The "scale," or scale factor, is denoted by $a(t)$, where $a(t)$ is measured in meters (or light-year, or Mpc, or whatever) per notch. The relation is then

$$\boxed{\ell_p(t) = a(t)\,\ell_c\ ,}$$

where $\ell_p(t)$ is the **physical** distance, measured in meters (or light-years, etc.), and ℓ_c is the **coordinate** distance, measured in notches.


z
x
y
z
x
y
z
x
y
"Notches"

This slide was not shown, but the work was done on the blackboard.

Hubble's Law as a Consequence of Uniform Expansion

$$\ell_p(t) = a(t)\,\ell_c \ ,$$

So how fast does $\ell_p(t)$ change?

$$v = \frac{\mathrm{d}\ell_p}{\mathrm{d}t} = \frac{\mathrm{d}a}{\mathrm{d}t}\ell_c = \left[\frac{1}{a(t)}\ \frac{\mathrm{d}a(t)}{\mathrm{d}t}\right] a(t)\ell_c \ .$$

Note that this can be rewritten as

$$v = \frac{\mathrm{d}\ell_p}{\mathrm{d}t} = H\ell_p \ ,$$ where $$H(t) = \frac{1}{a(t)}\ \frac{\mathrm{d}a(t)}{\mathrm{d}t} \ .$$

Light Rays in an Expanding Universe

☆ How do we describe light rays in the comoving coordinate system?

Light Rays in an Expanding Universe

- How do we describe light rays in the comoving coordinate system?
- The answer is simple:
  Light rays travel on a straight line, with a speed that would be measured by each local observer, as the light ray passes, at the standard value $c = 299,792,458$ m/s.
- Consider a light pulse moving along the x-axis. If the speed of light in m/s is c, and the number of meters per notch is $a(t)$, then the speed in notches per second is given by

$$\boxed{\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c}{a(t)}\ .}$$

- Justification: the above formula can be derived in general relativity by considering hypothetical point particles that travel at the speed of light, or by incorporating Maxwell's equations into general relativity.

This slide was not shown, but the work was done on the blackboard.

Importance of Comoving Coordinates

Any problem involving an expanding (homogeneous and isotropic) universe should be described in **comoving coordinates.**

Why?

Because the paths of light rays are simple in comoving coordinates.

If instead you tried to use coordinates that directly measure physical distances, the path of a light ray would be **complicated** for any trajectory other than a radial one.

Cosmic Time and the Synchronization of Clocks

☆ In special relativity, clocks can be synchronized by sending time signals from a central clock. Other clocks, when using these time signals, use their distances from the central clock to take into account the light travel time.

Cosmic Time and the Synchronization of Clocks

- ☆ In special relativity, clocks can be synchronized by sending time signals from a central clock. Other clocks, when using these time signals, use their distances from the central clock to take into account the light travel time.

- ☆ In an expanding universe, this does not work!

  1) Because the clocks are moving relative to each other, time dilation would have to be taken into account.

  2) Because the distances are changing with time, one can't know the distance until one knows the time, so the light travel time cannot be taken into account.

- In cosmology, we can imagine that "cosmic time" t is measured locally, on comoving clocks that tick in seconds defined by atomic standards. But they need to be synchronized somehow. Instead of using a central clock, one needs to find a clock that is available everywhere.

- In a simple, HOMOGENEOUS model of the universe, there are three possibilities:

  1) The Hubble expansion rate H. It can be measured anywhere, so can be used to define the $t = 0$ of cosmic time.

  2) The temperature T of the cosmic background radiation.

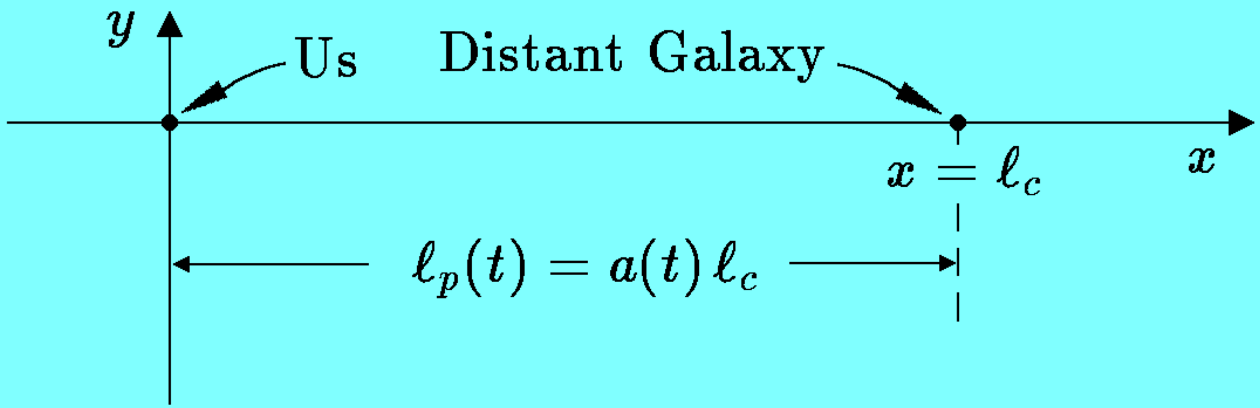
  3) If the universe starts with the scale factor $a = 0$, this starting time can be taken as $t = 0$.

- Will these three methods agree?

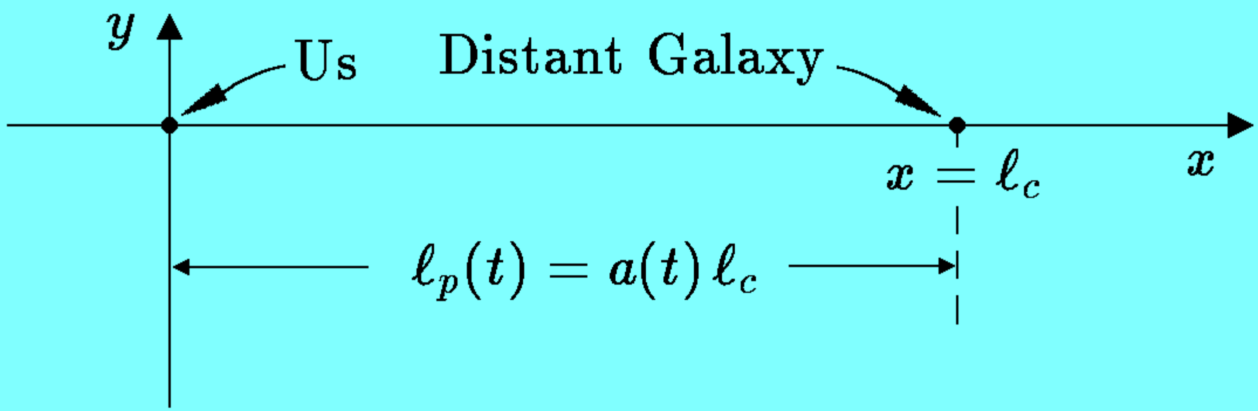
- Yes, they must, by the assumption of homogeneity. Homogeneity implies that the relation between H and T must be the same everywhere. So if you and I, in far away galaxies, measure the same value of H, we must also measure the same value of T.

Cosmological Redshift

- Use comoving coordinates!



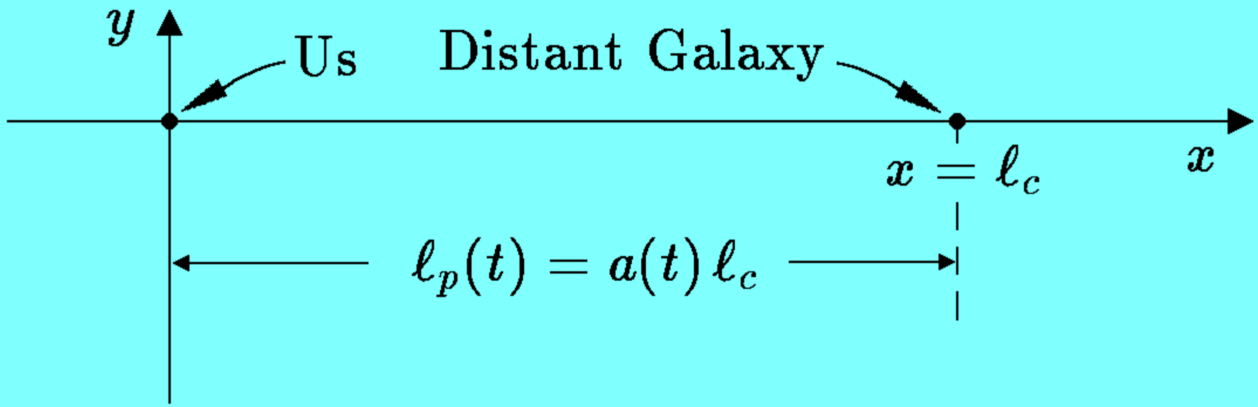
- Let Δt_S be the time between wave crests, as measured at the source.
- Since cosmic time t is measured on local clocks, Δt_S is the separation in cosmic time between the emission of crests.
- The physical wavelength at the source is $\lambda_S = c\Delta t_S$. When the 2nd crest is emitted, the first crest will be a physical distance λ_S from the source.
- When the 2nd crest is emitted, the first crest will be a coordinate distance $\Delta x = \lambda_S/a(t_S)$ from the source.



- When the 2nd crest is emitted, the first crest will be a coordinate distance $\Delta x = \lambda_S / a(t_S)$ from the source.
- As the first and second crests travel from source to us, they both travel at coordinate speed

$$\boxed{\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c}{a(t)}\ .}$$

- The speed depends on time, but not position: so the crests remain the same coordinate distance apart.



☆ When the crests reach us, at cosmic time t_O, they still have a coordinate separation $\Delta x = \lambda_S / a(t_S)$. The physical distance at the observer (us) is therefore

$$\boxed{\lambda_O = a(t_O)\,\Delta x = \left[\frac{a(t_O)}{a(t_S)}\right]\lambda_S \ ,}$$

so the wavelength is simply stretched with the expansion of the universe.

☆ The period of a light wave is proportional to its wavelength, so

$$\boxed{1 + z \equiv \frac{\Delta t_O}{\Delta t_S} = \frac{\lambda_O}{\lambda_S} = \frac{a(t_O)}{a(t_S)} \ .}$$

8.286 Lecture 5
September 21, 2022

THE DYNAMICS OF NEWTONIAN COSMOLOGY, PART 1

Isaac Newton to Richard Bentley, Letter 1
Newton on the Infinite Universe

As to your first query, it seems to me that if the matter of our sun and planets and all the matter of the universe were evenly scattered throughout all the heavens, and every particle had an innate gravity toward all the rest, and the whole space throughout which this matter was scattered was but finite, the matter on the outside of this space would, by its gravity, tend toward all the matter on the inside and, by consequence, fall down into the middle of the whole space and there compose one great spherical mass. But if the matter was evenly disposed throughout an infinite space, it could never convene into one mass; but some of it would convene into one mass and some into another, so as to make an infinite number of great masses, scattered at great distances from one to another throughout all that infinite space. And thus might the sun and fixed stars be formed, supposing the matter were of a lucid nature.

— December 10, 1692

But how the matter should divide itself into two sorts, and that part of it which is to compose a shining body should fall down into one mass and make a sun and the rest which is fit to compose an opaque body should coalesce, not into one great body, like the shining matter, but into many little ones; or if the sun at first were an opaque body like the planets or the planets lucid bodies like the sun, how he alone should be changed into a shining body whilst all they continue opaque, or all they be changed into opaque ones whilst he remains unchanged, I do not think explicable by mere natural causes, but am forced to ascribe it to the counsel and contrivance of a voluntary Agent.

— December 10, 1692

Web references: http://www.newtonproject.sussex.ac.uk/view/texts/normalized/THEM00254
http://books.google.com/books?id=8DkCAAAQAAJ&pg=PA201

Isaac Newton to Richard Bentley, Letter 2
Newton on Infinities

But you argue, in the next paragraph of your letter, that every particle of matter in an infinite space has an infinite quantity of matter on all sides, and, by consequence, an infinite attraction every way, and therefore must rest in equilibrio, because all infinites are equal. Yet you suspect a paralogism in this argument; and I conceive the paralogism lies in the position, that all infinites are equal. The generality of mankind consider infinites no other ways than indefinitely; and in this sense they say all infinites are equal; though they would speak more truly if they should say, they are neither equal nor unequal, nor have any certain difference or proportion one to another. In this sense, therefore, no conclusions can be drawn from them about the equality, proportions, or differences of things; and they that attempt to do it usually fall into paralogisms.

— January 17, 1693

So, when men argue against the infinite divisibility of magnitude, by saying, that if an inch may be divided into an infinite number of parts, the sum of those parts will be an inch; and if a foot may be divided into an infinite number of parts, the sum of those parts must be a foot; and therefore, since all infinites are equal, those sums must be equal, that is, an inch equal to a foot. The falseness of the conclusion shews an error in the premises ; and the error lies in the position, that all infinites are equal.

— January 17, 1693

Web references: http://www.newtonproject.sussex.ac.uk/view/texts/normalized/THEM00255
http://books.google.com/books?id=8DkCAAAAQAAJ&pg=PA201

Can a Uniform Infinite Distribution of Mass Be Stable?

Gauss's Law of Gravity:

$$\vec{g} = -\frac{GM}{r^2}\hat{r} \quad \Longrightarrow \quad \boxed{\oint \vec{g} \cdot d\vec{a} = -4\pi G M_{\text{enclosed}}}$$

Poisson's Equation:

$$\boxed{\nabla^2\phi = 4\pi G\rho\ ,}\quad \text{where } \vec{g} = -\vec{\nabla}\phi\ .$$

where ρ is the mass density, $\vec{\nabla}\phi$ is the gradient of ϕ:

$$\vec{\nabla}\phi \equiv \hat{\imath}\,\frac{\partial\phi}{\partial x} + \hat{\jmath}\frac{\partial\phi}{\partial y} + \hat{k}\frac{\partial\phi}{\partial z}\ ,$$

and $\nabla^2\phi$ is the Laplacian of ϕ:

$$\nabla^2\phi \equiv \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}\ .$$

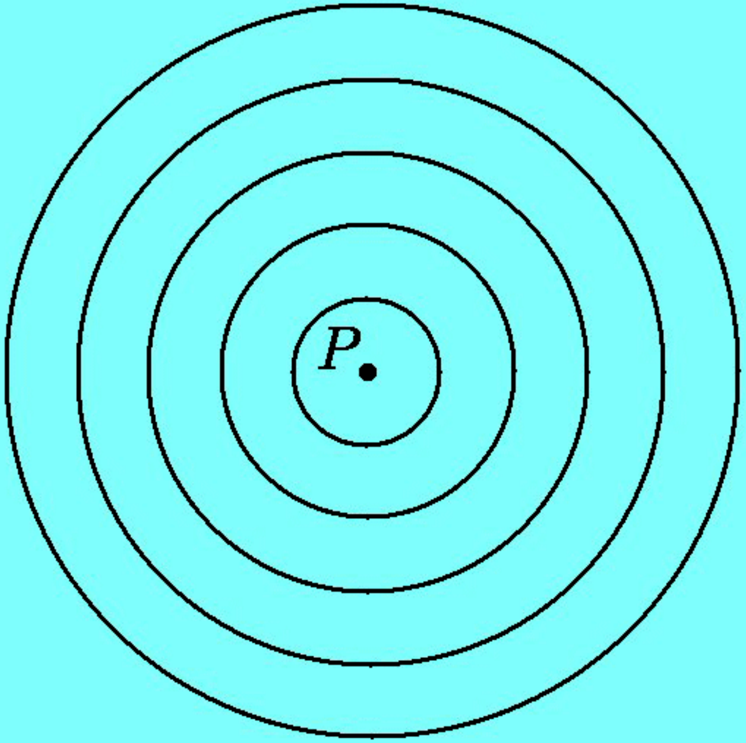
and

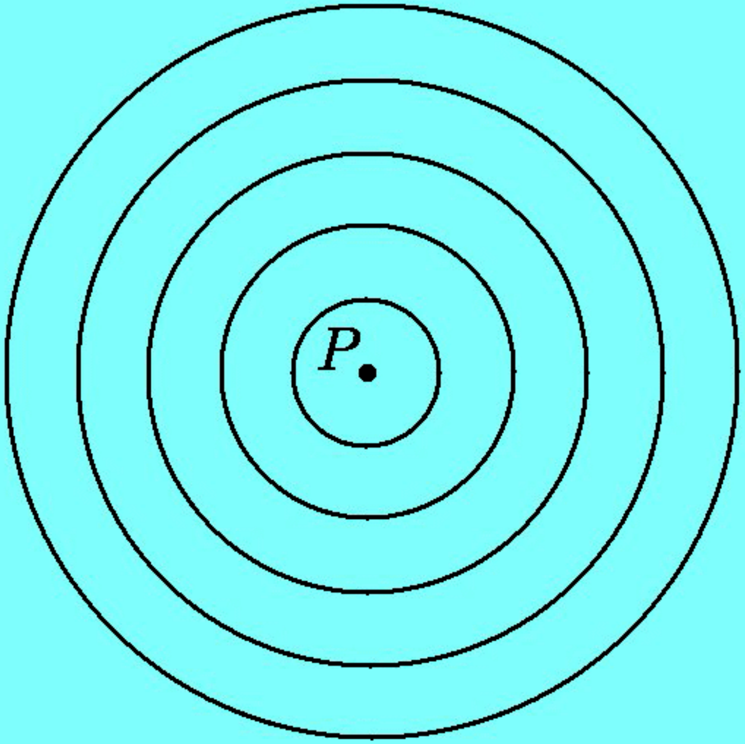
Attempting to Integrate the Gravitational Force

☆ Problem: the integral is not *absolutely* convergent, but only *conditionally* convergent.

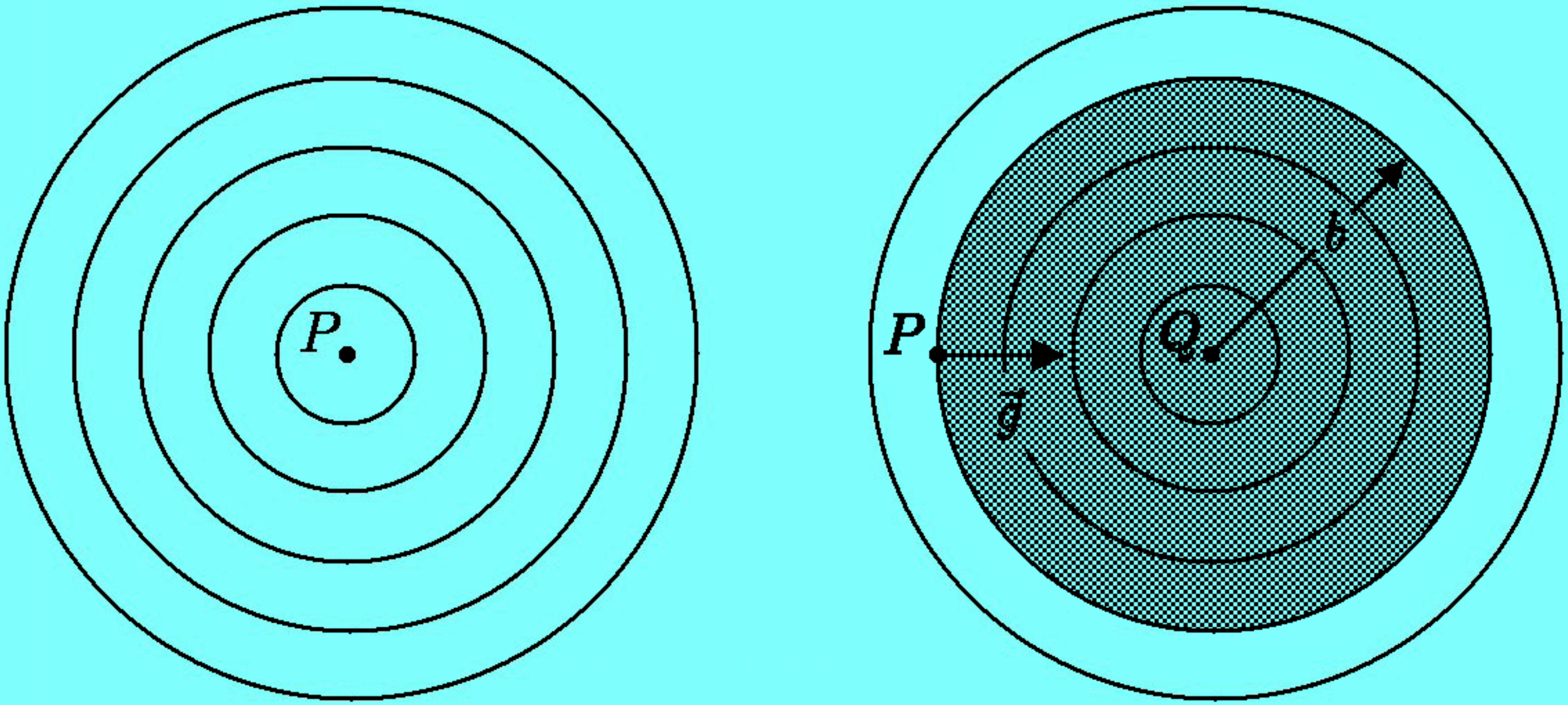
Attempting to Integrate the Gravitational Force

- Problem: the integral is not *absolutely* convergent, but only *conditionally* convergent.
- Discussion of absolute and conditional convergence (on blackboard).

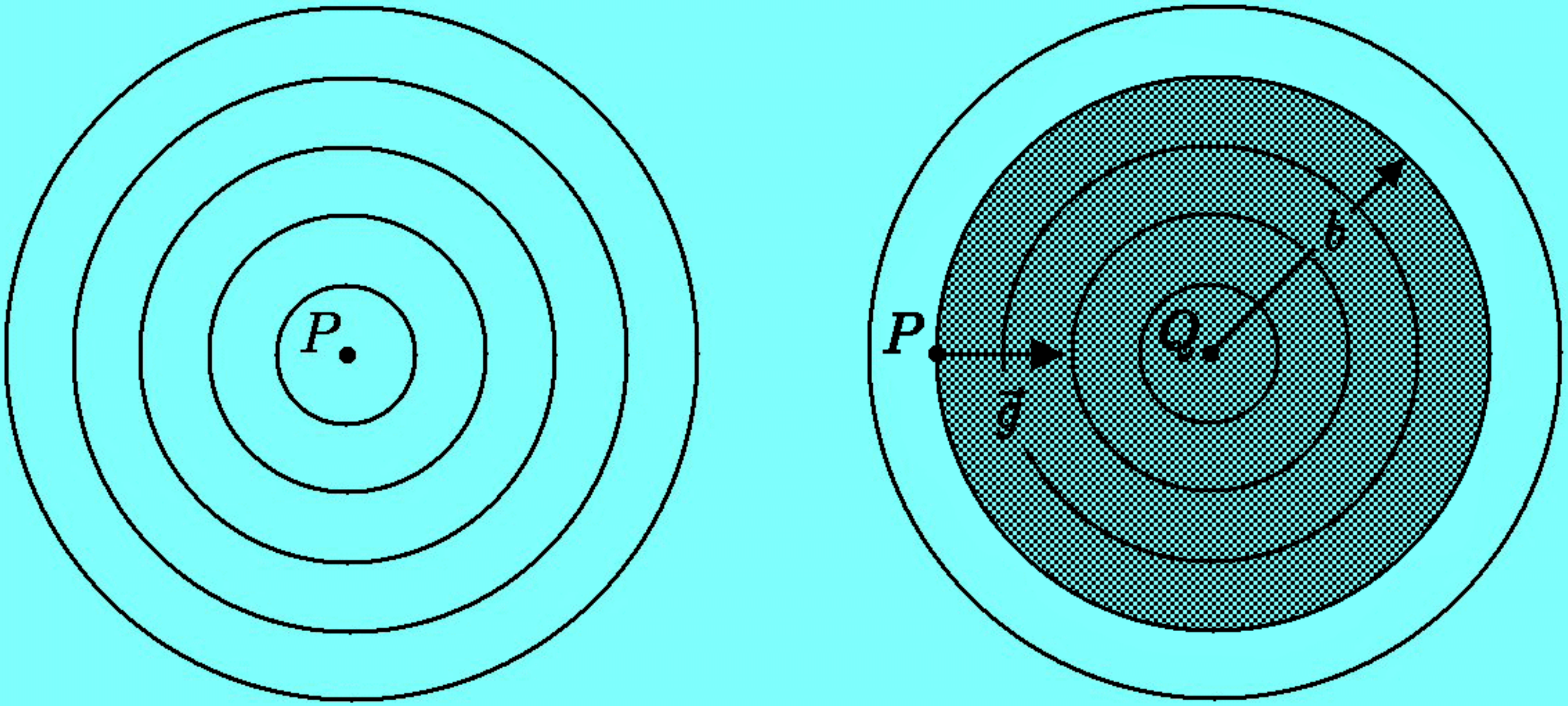

P



$\vec{g} = 0$



$\vec{g} = 0$



$$\vec{g} = 0$$

$$\vec{g} = \frac{GM}{b^2}\hat{e}_{QP}$$

But What About Symmetry?

- ☆ Newton argued that there could be no acceleration, because there is no preferred direction for it to point.

But What About Symmetry?

- Newton argued that there could be no acceleration, because there is no preferred direction for it to point.
- Complication: acceleration is measured relative to an inertial frame, which Newton defined as the frame of the "fixed stars". But if the universe collapses, then there are no fixed stars.

But What About Symmetry?

- Newton argued that there could be no acceleration, because there is no preferred direction for it to point.
- Complication: acceleration is measured relative to an inertial frame, which Newton defined as the frame of the "fixed stars". But if the universe collapses, then there are no fixed stars.
- In the absence of an inertial frame, all accelerations, like velocities, are relative.

But What About Symmetry?

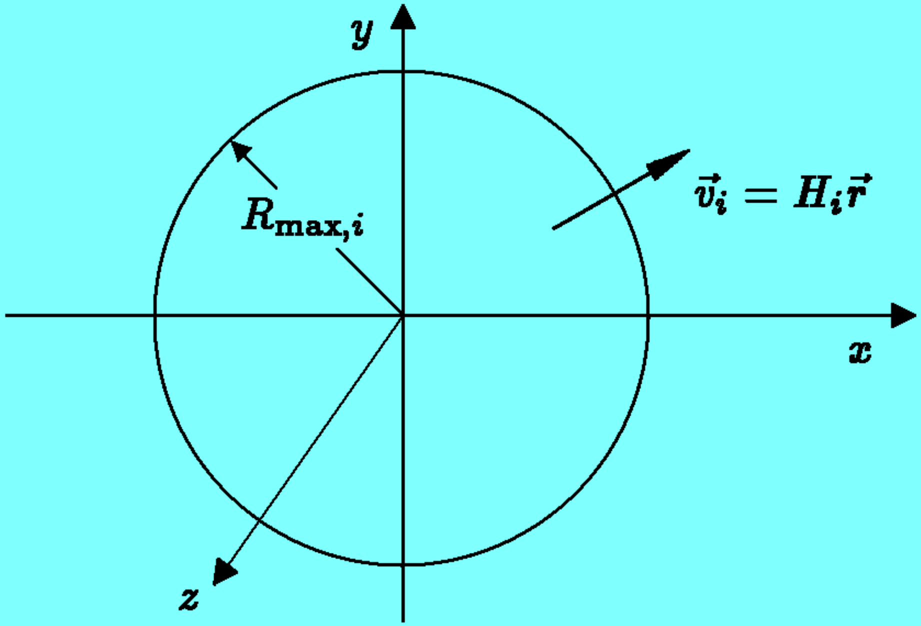
- Newton argued that there could be no acceleration, because there is no preferred direction for it to point.
- Complication: acceleration is measured relative to an inertial frame, which Newton defined as the frame of the "fixed stars". But if the universe collapses, then there are no fixed stars.
- In the absence of an inertial frame, all accelerations, like velocities, are relative.
- When all accelerations are relative, any observer can consider herself to be non-accelerating. She would then see all other objects accelerating radially toward herself. Like the velocities of Hubble expansion, this picture looks like it has a unique center, but really it is homogenous.

Mathematical Model of a Uniformly Expanding Universe

- Desired properties: homogeneity, isotropy, and Hubble's law.
- The model should be finite, to avoid the conditional convergence problems discussed last time. At the end we will take the limit as the size approaches infinity.
- Newtonian dynamics: we choose the initial conditions, and then Newton's laws of motion will determine how it will evolve.
- To impose isotropy, we model the initial state as a solid sphere, of some radius $R_{\mathrm{max},i}$.
- To impose homogeneity, we take the initial mass density to be constant, ρ_i. The matter is treated as a dust, that can thin as the universe expands. By dust, we mean a system of very small closely spaced particles, that can be treated as a continuous fluid, with negligible pressure.
- We take the initial velocities according to Hubble's law, with some initial expansion rate H_i

Epochs of Cosmic History

Mathematical Model of a Uniformly Expanding Universe



$$t_i \equiv \text{time of initial picture}$$

$$R_{\max,i} \equiv \text{initial maximum radius}$$

$$\rho_i \equiv \text{initial mass density}$$

$$\vec{v}_i = H_i \vec{r}\ .$$

Description of Evolution

- As the model universe evolves, the spherical symmetry will be preserved: each gas particle will continue on a radial trajectory, since there are no forces that might pull it tangentially.

- Spherical symmetry $\Longrightarrow$ all particles that start at the same initial radius will behave the same way. So, a particle that begins at radius r_i will be found at a later time t at some radius

$$\boxed{r = r(r_i, t)\ .}$$

- Our goal is to figure out what determines $r(r_i, t)$

- The only relevant force is gravity. Gravity and electromagnetism are the only (known) long-range forces. The universe appears to be electrically neutral, so long-range electric forces are not present.

Reminder: the Gravitational Field of a Shell of Matter

- ☆ For points outside the shell, the gravitational force is the same as if the total mass of the shell were concentrated at the center.
- ☆ For points inside the shell, the gravitational field is **zero.**
- ☆ Newton figured this out by integration. For us, Gauss's law makes it obvious.

Shell Crossings?

Can shells cross? I.e., can two shells that start at different r_i ever cross each other?

The answer is no, but we don't know that when we start.

But we do know that Hubble's law implies that any two shells are initially moving apart. Therefore there must be at least some interval before any shell crossings can happen.

We will write equations that are valid assuming no shell crossings.

These equations will be valid until any possible shell crossing.

If there was a shell crossing, these equations would have to show two shells becoming arbitrarily close.

We will find, however, that the equations imply uniform expansion, so no shell crossings ever happen in this system.

Equations of Motion

- Newtonian gravity of a shell:

  Inside: $\vec{g} = 0$.

  Outside: Same as point mass at center, with same M.

- $r(r_i, t) \equiv$ radius at t of shell initially at r_i.

- Let $M(r_i) \equiv$ mass inside r_i-shell $= \frac{4\pi}{3} r_i^3 \rho_i$ at all times.

- Pressure? When a gas with pressure $p > 0$ expands, it pushes on its surroundings and loses energy. Relativistically, energy = mass (times c^2). By assuming that $M(r_i)$ is constant, we are assuming that $p \simeq 0$.

Equations:

☆ For particles at radius r,

$$\boxed{\vec{g} = -\frac{GM(r_i)}{r^2}\,\hat{r}\ ,}$$

where

$$\boxed{M(r_i) = \frac{4\pi}{3} r_i^3 \rho_i\ .}$$

Since $\vec{g}$ is the acceleration,

$$\boxed{\ddot{r} = -\frac{GM(r_i)}{r^2} = -\frac{4\pi}{3}\frac{G r_i^3 \rho_i}{r^2}\ , \quad \text{where } r \equiv r(r_i, t),}$$

where an overdot indicates a derivative with respect to t.

$$\ddot{r} = -\frac{GM(r_i)}{r^2} = -\frac{4\pi}{3}\frac{Gr_i^3\rho_i}{r^2} \ , \ \text{where } r \equiv r(r_i, t),$$

☆ For a second order equation like this, the solution is uniquely determined if the initial value of r and $\dot{r}$ are specified:

$$r(r_i, t_i) = r_i \ ,$$

and, by the Hubble law initial condition $\vec{v}_i = H_i \vec{r}_i$,

$$\dot{r}(r_i, t_i) = H_i r_i \ .$$

Miraculous Scaling Relations

$$\ddot{r} = -\frac{4\pi}{3}\frac{G r_i^3 \rho_i}{r^2} \ , \quad r(r_i, t_i) = r_i \ , \quad \dot{r}(r_i, t_i) = H_i r_i \ .$$

⭐ Suppose we define

$$u(r_i, t) \equiv \frac{r(r_i, t)}{r_i} \ .$$

Then

$$\ddot{u} = \frac{\ddot{r}}{r_i} = -\frac{4\pi}{3}\frac{G\rho_i}{u^2} \ .$$

There is no r_i-dependence. This "miracle" depended on gravity being a $1/r^2$ force.

$$\ddot{r} = -\frac{4\pi}{3}\frac{G r_i^3 \rho_i}{r^2} \;, \quad r(r_i, t_i) = r_i \;, \quad \dot{r}(r_i, t_i) = H_i r_i \;.$$

$$u(r_i, t) \equiv \frac{r(r_i, t)}{r_i} \quad \Longrightarrow \quad \ddot{u} = -\frac{4\pi}{3}\frac{G\rho_i}{u^2} \;.$$

What about the initial conditions for $u(r_i, t)$?

$$u(r_i, t_i) = \frac{r(r_i, t_i)}{r_i} = 1 \;, \quad \dot{u}(r_i, t_i) = \frac{\dot{r}(r_i, t_i)}{r_i} = H_i \;.$$

Since the differential equation and the intial conditions determine $u(r_i, t)$, it does not depend on r_i. We can rename it

$$u(r_i, t) \equiv a(t) \;, \qquad \text{so} \qquad r(r_i, t) = a(t)\, r_i \;.$$

This describes uniform expansion by a scale factor $a(t)$.

8.286 Lecture 6
September 26, 2022

THE DYNAMICS OF NEWTONIAN COSMOLOGY, PART 2

Can a Uniform Infinite Distribution of Mass Be Stable?

- Newton (1692): Yes.
- Gauss's Law of Gravity (Lagrange, 1773, Gauss, 1835): No, but not noticed by anybody.

$$\vec{g} = -\frac{GM}{r^2}\hat{r} \quad \Longrightarrow \quad \boxed{\oint \vec{g} \cdot d\vec{a} = -4\pi G M_{\text{enclosed}}}$$

- Poisson's Equation (1829): No, but not noticed by anyone.

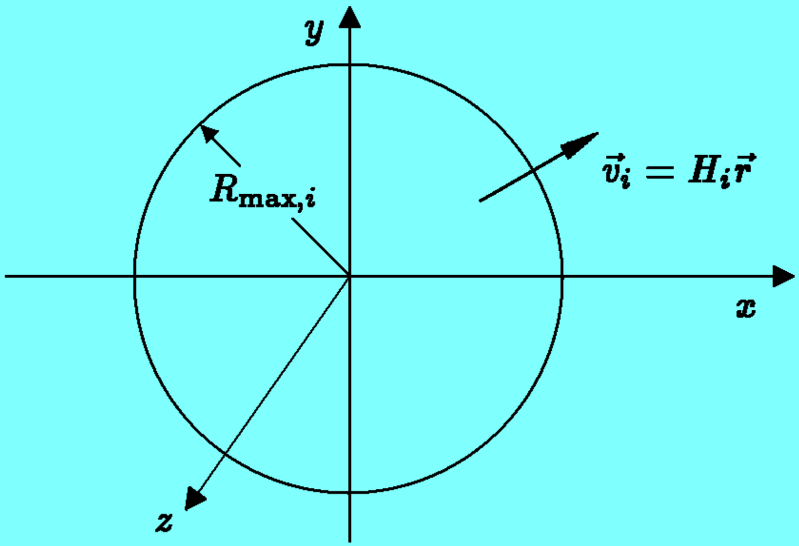
$$\boxed{\nabla^2 \phi = 4\pi G \rho\ ,} \quad \text{where } \vec{g} = -\vec{\nabla}\phi\ ,$$

where ρ is the mass density.

☆ Einstein (1917): No.

Einstein discovered that a uniform infinite distribution of mass is unstable in Newtonian physics, and also in the original form of General Relativity. (Einstein was, however, still convinced that the universe must be static. He found that he could add a new term to the equations that describe how matter creates a gravitational field (the Einstein field equations), which he called the *cosmological term*, which described a repulsive force which could be adjusted to just balance the attractive force, allowing a static universe. Einstein endorsed this model until Hubble discovered in 1929 that the universe was expanding.

Mathematical Model



$$t_i \equiv \text{time of initial picture}$$

$$R_{\max,i} \equiv \text{initial maximum radius}$$

$$\rho_i \equiv \text{initial mass density}$$

$$\vec{v}_i = H_i\vec{r}\ .$$

Miraculous Scaling Relations

$$\ddot{r} = -\frac{4\pi}{3}\frac{G r_i^3 \rho_i}{r^2} \ , \quad r(r_i, t_i) = r_i \ , \quad \dot{r}(r_i, t_i) = H_i r_i \ .$$

☆ Suppose we define

$$u(r_i, t) \equiv \frac{r(r_i, t)}{r_i} \ .$$

Then

$$\ddot{u} = \frac{\ddot{r}}{r_i} = -\frac{4\pi}{3}\frac{G\rho_i}{u^2} \ .$$

There is no r_i-dependence. This "miracle" depended on gravity being a $1/r^2$ force.

$$\ddot{r} = -\frac{4\pi}{3}\frac{G r_i^3 \rho_i}{r^2} \ , \quad r(r_i, t_i) = r_i \ , \quad \dot{r}(r_i, t_i) = H_i r_i \ .$$

$$u(r_i, t) \equiv \frac{r(r_i, t)}{r_i} \quad \Longrightarrow \quad \ddot{u} = -\frac{4\pi}{3}\frac{G\rho_i}{u^2} \ .$$

What about the initial conditions for $u(r_i, t)$?

$$u(r_i, t_i) = \frac{r(r_i, t_i)}{r_i} = 1 \ , \quad \dot{u}(r_i, t_i) = \frac{\dot{r}(r_i, t_i)}{r_i} = H_i \ .$$

Since the differential equation and the intial conditions determine $u(r_i, t)$, it does not depend on r_i. We can rename it

$$\boxed{u(r_i, t) \equiv a(t) \ ,} \quad \text{so} \quad \boxed{r(r_i, t) = a(t)\, r_i \ .}$$

This describes uniform expansion by a scale factor $a(t)$.

This slide was not shown, but the work was done on the blackboard.

Time Dependence of $\rho(t)$

We know how the mass density depends on time, because we assumed that $M(r_i)$ — the total mass contained inside a shell of particles whose initial radius was r_i — does not change with time. The radius of the shell at time t is $a(t)r_i$. The mass density is just the mass divided by the volume,

$$\boxed{\rho(t) = \frac{M(r_i)}{\frac{4\pi}{3}a^3(t)r_i^3} = \frac{\frac{4\pi}{3}r_i^3\rho_i}{\frac{4\pi}{3}a^3(t)r_i^3} = \frac{\rho_i}{a^3(t)} \ .}$$

So

$$\boxed{\ddot{u} = -\frac{4\pi}{3}\frac{G\rho_i}{u^2} \quad \Longrightarrow \quad \ddot{a} = -\frac{4\pi}{3}\frac{G\rho_i}{a^2} \ .}$$

$$\Longrightarrow \quad \boxed{\ddot{a} = -\frac{4\pi}{3}G\rho(t)\,a(t) \ .} \quad \text{Friedmann equation.}$$

Nothing Depends on $R_{\text{max},i}$

- An observer living in this model universe would see uniform expansion all around herself, and would only be aware of the boundary at R_{max} if she was close enough to the boundary to see it.
- Thus, we can take the limit $R_{\text{max},i} \to \infty$ without doing anything, since nothing of interest depends on $R_{\text{max},i}$.

This slide was not shown, but the work was done on the blackboard.

A Conservation Law

☆ The equation for $\ddot{a}$ has the same form as an equation for the motion of a particle with a time-independent potential energy function. So, there is a conservation law:

$$\ddot{a} = -\frac{4\pi}{3}\frac{G\rho_i}{a^2} \quad \Longrightarrow \quad \dot{a}\left\{\ddot{a} + \frac{4\pi}{3}\frac{G\rho_i}{a^2}\right\} = 0 \quad \Longrightarrow \quad \frac{dE}{dt} = 0 ,$$

where

$$E = \frac{1}{2}\dot{a}^2 - \frac{4\pi}{3}\frac{G\rho_i}{a} .$$

Summary: Equations

Want: $r(r_i, t) \equiv$ radius at t of shell initially at r_i

Find: $r(r_i, t) = a(t) r_i$, where

$$\text{Friedmann Equations} \begin{cases} \ddot{a} = -\dfrac{4\pi}{3} G \rho(t) a \\ H^2 = \left(\dfrac{\dot{a}}{a}\right)^2 = \dfrac{8\pi}{3} G\rho - \dfrac{kc^2}{a^2} \quad \text{(Friedmann Eq.)} \end{cases}$$

and

$$\rho(t) \propto \frac{1}{a^3(t)} \ , \ \text{or } \rho(t) = \left[\frac{a(t_1)}{a(t)}\right]^3 \rho(t_1) \text{ for any } t_1 .$$

☆ Note that t_i no longer plays any role. It does not appear on this slide!

The Return of the 'Notch'

- Definition: $r(r_i, t) = a(t) r_i$.
- In the previous derivation, r_i was the initial radius of some particle, measured in meters. But when we finished, r_i was being used only as a coordinate to label shells, where the shell corresponding to $r_i = 1$ had a radius of one meter only at time t_i.
- But t_i no longer appears, and will not be mentioned again! So, the connection between the numerical value of r_i and the length of a meter has disappeared from the formalism.
- Bottom line: r_i is the radial coordinate in a comoving coordinate system, measured in units that have no particular meaning. I will refer to the units of r_i as "notches," but you should be aware that the term is not standard.

This slide was not shown, but the work was done on the blackboard.

Conventions for the Notch

Us: For us, the notch is an arbitrary unit that we use to mark off intervals on the comoving coordinate system. We are free to use a different definition every time we use the notch.

Ryden: $a(t_0) = 1$ (where $t_0 =$ now). (In our language, Ryden's convention is $a(t_0) = 1$ m/notch.)

Many Other Books: if $k \neq 0$, then $k = \pm 1$.

In our language, this means $k = \pm 1/\text{notch}^2$. To see the units of k, recall that the Friedmann equation is

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3} G\rho - \frac{kc^2}{a^2} \ .$$

We will use $[x]$ to mean the units of x, and we will use T and L to denote the units of time and length, respectively. The units of the left-hand side are $1/T^2$, with the units of a canceling. So

$$[k] = \frac{1}{T^2}\left[\frac{a}{c}\right]^2 = \frac{1}{T^2}\left[\frac{L/\text{notch}}{L/T}\right]^2 = \frac{1}{\text{notch}^2} \ .$$

This slide was not shown, but the work was done on the blackboard.

Types of Solutions

$$\dot{a}^2 = \frac{8\pi G}{3} \frac{\rho(t_1)a^3(t_1)}{a(t)} - kc^2 \quad \text{(for any } t_1\text{)} \ .$$

For intuition, remember that $k \propto -E$, where E is a measure of the energy of the system.

Types of Solutions:

1) $k < 0$ $(E > 0)$: unbound system. $\dot{a}^2 > (-kc^2) > 0$, so the universe expands forever. **Open Universe.**

2) $k > 0$ $(E < 0)$: bound system. $\dot{a}^2 \geq 0 \quad \Longrightarrow$

$$a_{\mathrm{max}} = \frac{8\pi G}{3} \frac{\rho(t_1)a^3(t_1)}{kc^2} \ .$$

Universe reaches maximum size and then contracts to a Big Crunch. **Closed Universe.**

3) $k = 0$ ($E = 0$): critical mass density.

$$H^2 = \frac{8\pi G}{3}\rho - \underbrace{\frac{kc^2}{a^2}}_{=0} \quad \Longrightarrow \quad \boxed{\rho \equiv \rho_c = \frac{3H^2}{8\pi G}\ .}$$

Flat Universe.

Summary: $\rho > \rho_c \Longleftrightarrow$ closed, $\rho < \rho_c \Longleftrightarrow$ open, $\rho = \rho_c \Longleftrightarrow$ flat.

Numerical value: For $H = 68$ km-s^{-1}-Mpc^{-1} (Planck 2015 plus other experiments),

$$\rho_c = 8.7 \times 10^{-27}\ \mathrm{kg/m^3} = 8.7 \times 10^{-30}\ \mathrm{g/cm^3}$$

$$\approx 5 \text{ proton masses per } \mathrm{m}^3.$$

Definition: $\Omega \equiv \dfrac{\rho}{\rho_c}$.

8.286 Lecture 7
September 28, 2022

THE DYNAMICS OF NEWTONIAN COSMOLOGY, PART 3

Review from the previous lecture

Equations for a Matter-Dominated Universe

("Matter-dominated" = dominated by nonrelativistic matter.)

Friedmann equations:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} \; ,$$

$$\ddot{a} = -\frac{4\pi}{3}G\rho(t)a \; .$$

Matter conservation:

$$\rho(t) \propto \frac{1}{a^3(t)} \; , \text{ or } \rho(t) = \left[\frac{a(t_1)}{a(t)}\right]^3 \rho(t_1) \text{ for any } t_1 .$$

Any two of the above equations can allow us to find the third.

To make contact with our earlier notation,

$$k = -\frac{2E}{c^2} \ ,$$

where

$$E = \frac{1}{2}\dot{a}^2 - \frac{4\pi}{3}\frac{G\rho_i}{a} \ .$$

We discovered that E is conserved, as a consequence of the second order ($\ddot{a}$) equation of motion.

Review of Comoving Coordinates

☆ Original form of our Newtonian derivation:

$$\begin{aligned} t_i &= \text{initial time} \\ r_i &= \text{initial radius of some given particle} \\ r(r_i, t) &= a(t) r_i \ . \end{aligned}$$

r_i would be measured in physical length units, e.g. meters.

☆ Deprecating t_i:

We could have started the same model at any time. t_i is not really a property of the model universe. So

$$r_i = \text{radius of the given particle in meters at time } t_i$$

☆ Redefinitions to eliminate t_i:

Use "$\longrightarrow$" to mean "rename".

1 m at time $t_i \longrightarrow$ "notch"

$r_i \longrightarrow r_c$ (coordinate radius, measured in notches)

Then the physical radius r_p of any particle, at any time t, is given by

$$\boxed{r_p(t) = a(t) r_c \ .}$$

Types of Solutions

$$\dot{a}^2 = \frac{8\pi G}{3} \frac{\rho(t_1) a^3(t_1)}{a(t)} - kc^2 \quad (\text{for any } t_1) \; .$$

For intuition, remember that $k \propto -E$, where E is a measure of the energy of the system.

Types of Solutions:

1) $k < 0$ $(E > 0)$: unbound system. $\dot{a}^2 > (-kc^2) > 0$, so the universe expands forever. **Open Universe.**

2) $k > 0$ $(E < 0)$: bound system. $\dot{a}^2 \geq 0 \quad \Longrightarrow$

$$a_{\max} = \frac{8\pi G}{3} \frac{\rho(t_1) a^3(t_1)}{kc^2} \; .$$

Universe reaches maximum size and then contracts to a Big Crunch. **Closed Universe.**

3) $k = 0$ $(E = 0)$: critical mass density.

$$H^2 = \frac{8\pi G}{3}\rho - \underbrace{\frac{kc^2}{a^2}}_{=0} \quad \Longrightarrow \quad \boxed{\rho \equiv \rho_c = \frac{3H^2}{8\pi G}\ .}$$

Flat Universe.

Summary: $\rho > \rho_c \Longleftrightarrow$ closed, $\rho < \rho_c \Longleftrightarrow$ open, $\rho = \rho_c \Longleftrightarrow$ flat.

Numerical value: For $H = 68$ km-s^{-1}-Mpc^{-1} (Planck 2015 plus other experiments),

$$\rho_c = 8.7 \times 10^{-27}\ \mathrm{kg/m^3} = 8.7 \times 10^{-30}\ \mathrm{g/cm^3}$$

$$\approx 5 \text{ proton masses per m}^3.$$

Definition: $\Omega \equiv \dfrac{\rho}{\rho_c}$.

Evolution of a Flat Matter-Dominated Universe

If $k = 0$, then

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho = \frac{\text{const}}{a^3} \quad \Longrightarrow \quad \frac{da}{dt} = \frac{\text{const}}{a^{1/2}}$$

$$\Longrightarrow \quad a^{1/2}\, da = \text{const}\, dt \quad \Longrightarrow \quad \frac{2}{3}a^{3/2} = (\text{const})t + c' \ .$$

Choose the zero of time to make $c' = 0$, and then

$$\boxed{a(t) \propto t^{2/3} \ .}$$

Age of a Flat Matter-Dominated Universe

$$a(t) \propto t^{2/3} \quad \Longrightarrow \quad H = \frac{\dot{a}}{a} = \frac{2}{3t} \quad \Longrightarrow$$

$$\boxed{t = \frac{2}{3} H^{-1}}$$

For $H = 67.7 \pm 0.5$ km-s^{-1}-Mpc^{-1}, age $= 9.56$ – 9.70 billion years — but stars are older. Conclusion: our universe is nearly flat, but not matter-dominated.

Big Bang Singularity

- $a(0) = 0$, so the mass density ρ at $t = 0$ is infinite.
- This instant of infinite mass density is called a singularity.
- But, as we extrapolate backwards to early t, ρ becomes higher than any mass density that we know about.
- Hence, there is no reason to trust the model back to $t = 0$.

Big Bang Singularity

- ☆ $a(0) = 0$, so the mass density ρ at $t = 0$ is infinite.
- ☆ This instant of infinite mass density is called a singularity.
- ☆ But, as we extrapolate backwards to early t, ρ becomes higher than any mass density that we know about.
- ☆ Hence, there is no reason to trust the model back to $t = 0$.
- ☆ Conclusion: the singularity is a feature of the model, but not necessarily the real universe.

Big Bang Singularity

- $a(0) = 0$, so the mass density ρ at $t = 0$ is infinite.
- This instant of infinite mass density is called a singularity.
- But, as we extrapolate backwards to early t, ρ becomes higher than any mass density that we know about.
- Hence, there is no reason to trust the model back to $t = 0$.
- Conclusion: the singularity is a feature of the model, but not necessarily the real universe.
- Quantum gravity? The singularity is a feature of the *classical* theory, but might be avoided by a quantum gravity treatment — but we don't know.

Big Bang Singularity

- $a(0) = 0$, so the mass density ρ at $t = 0$ is infinite.
- This instant of infinite mass density is called a singularity.
- But, as we extrapolate backwards to early t, ρ becomes higher than any mass density that we know about.
- Hence, there is no reason to trust the model back to $t = 0$.
- Conclusion: the singularity is a feature of the model, but not necessarily the real universe.
- Quantum gravity? The singularity is a feature of the *classical* theory, but might be avoided by a quantum gravity treatment — but we don't know.
- In eternal inflation models, to be discussed near the end of the term, the event 13.8 billion years ago was not a singularity, but rather the decay of the repulsive-gravity material that drove the inflation. There might still have been a singularity deeper in the past.

This slide was not shown, but the work was done on the blackboard.

Horizon Distance

Definition: the horizon distance is the present distance of the furthest particles from which light has had time to reach us.

To find it, use comoving coordinates. The coordinate velocity of light is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c}{a(t)} \ ,$$

so the maximum coordinate distance that light could have traveled by time t (starting at $t = 0$) is

$$\ell_{c,\mathrm{horizon}}(t) = \int_0^t \frac{c}{a(t')}\,\mathrm{d}t' \ .$$

$$\ell_{c,\text{horizon}}(t) = \int_0^t \frac{c}{a(t')}\,\mathrm{d}t' \ .$$

The horizon distance is the maximum *physical* distance that light could have traveled, so

$$\boxed{\ell_{\text{phys,horizon}}(t) = a(t)\int_0^t \frac{c}{a(t')}\mathrm{d}t' \ .}$$

For a flat, matter-dominated universe, $a(t) \propto t^{2/3}$, so

$$\boxed{\ell_{\text{phys,horizon}}(t) = 3ct = 2cH^{-1} \ ,}$$

since $t = \frac{2}{3}H^{-1}$.

Evolution of a Closed Matter-Dominated Universe

$$\left(\frac{\dot a}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}\ , \quad \rho(t)a^3(t) = \text{constant}\ ,\ k > 0\ .$$

Recall $[a(t)] = \text{meter}/\text{notch}$, $[k] = 1/\text{notch}^2$.

Define new variables:

$$\tilde a(t) \equiv \frac{a(t)}{\sqrt{k}}\ , \qquad \tilde t \equiv ct \qquad \text{(both with units of distance)}$$

Multiplying Friedmann eq by $a^2/(kc^2)$:

$$\frac{1}{kc^2}\left(\frac{da}{dt}\right)^2 = \frac{8\pi}{3}\frac{G\rho a^2}{kc^2} - 1\ .$$

Recalling

$$\tilde{a}(t) \equiv \frac{a(t)}{\sqrt{k}} \, , \qquad \tilde{t} \equiv ct,$$

we find

$$\frac{1}{kc^2}\left(\frac{da}{dt}\right)^2 = \frac{8\pi}{3}\frac{G\rho a^2}{kc^2} - 1$$

$$= \frac{8\pi}{3}\frac{G\rho a^3}{k^{3/2}c^2}\,\frac{\sqrt{k}}{a} - 1 \, .$$

Rewrite as

$$\left(\frac{d\tilde{a}}{d\tilde{t}}\right)^2 = \frac{2\alpha}{\tilde{a}} - 1 \, ,$$

where

$$\alpha \equiv \frac{4\pi}{3}\frac{G\rho\tilde{a}^3}{c^2} \, .$$

$[\alpha]$ = meter. α is constant, since ρa^3 is constant.

$$\left(\frac{d\tilde{a}}{d\tilde{t}}\right)^2 = \frac{2\alpha}{\tilde{a}} - 1 \quad \Longrightarrow \quad d\tilde{t} = \frac{\tilde{a}\, d\tilde{a}}{\sqrt{2\alpha\tilde{a} - \tilde{a}^2}} \ .$$

Then

$$\tilde{t}_f = \int_0^{\tilde{t}_f} d\tilde{t} = \int_0^{\tilde{a}_f} \frac{\tilde{a}\, d\tilde{a}}{\sqrt{2\alpha\tilde{a} - \tilde{a}^2}} \ ,$$

where $\tilde{t}_f$ is an arbitrary choice for a "final time" for the calculation, and $\tilde{a}_f$ is the value of $\tilde{a}$ at time $\tilde{t}_f$.

To carry out the integral, we first complete the square:

$$\tilde{t}_f = \int_0^{\tilde{a}_f} \frac{\tilde{a}\, d\tilde{a}}{\sqrt{\alpha^2 - (\tilde{a} - \alpha)^2}} \ .$$

Now simplify by defining $x \equiv \tilde{a} - \alpha$, so

$$\tilde{t}_f = \int_{-\alpha}^{\tilde{a}_f - \alpha} \frac{(x + \alpha)\, dx}{\sqrt{\alpha^2 - x^2}} \ .$$

$$\tilde{t}_f = \int_{-\alpha}^{\tilde{a}_f - \alpha} \frac{(x+\alpha)\,dx}{\sqrt{\alpha^2 - x^2}} \ .$$

To simplify $\alpha^2 - x^2$, define θ so that $x = -\alpha\cos\theta$.

(Choice of the minus sign simplifies the final answer. Recall that x represents the scale factor, and θ will be replacing x. The minus sign leads to $\mathrm{d}x/\mathrm{d}\theta = \alpha\sin\theta$, which is positive for small positive θ, so both will be growing at the start of the universe.)

Substituting,

$$\mathrm{d}x = \alpha\sin\theta\,\mathrm{d}\theta \ .$$

$$\sqrt{\alpha^2 - x^2} = \alpha\sqrt{1-\cos^2\theta} = \alpha\sin\theta \ .$$

Then

$$\tilde{t}_f = \alpha\int_0^{\theta_f}(1-\cos\theta)d\theta = \alpha(\theta_f - \sin\theta_f) \ .$$

This equation relates t_f to θ_f, but we really want to relate the scale factor and time. But θ_f is related to the scale factor, if we trace back the definitions: $x_f = -\alpha\cos\theta_f = \tilde{a}_f - \alpha$, so

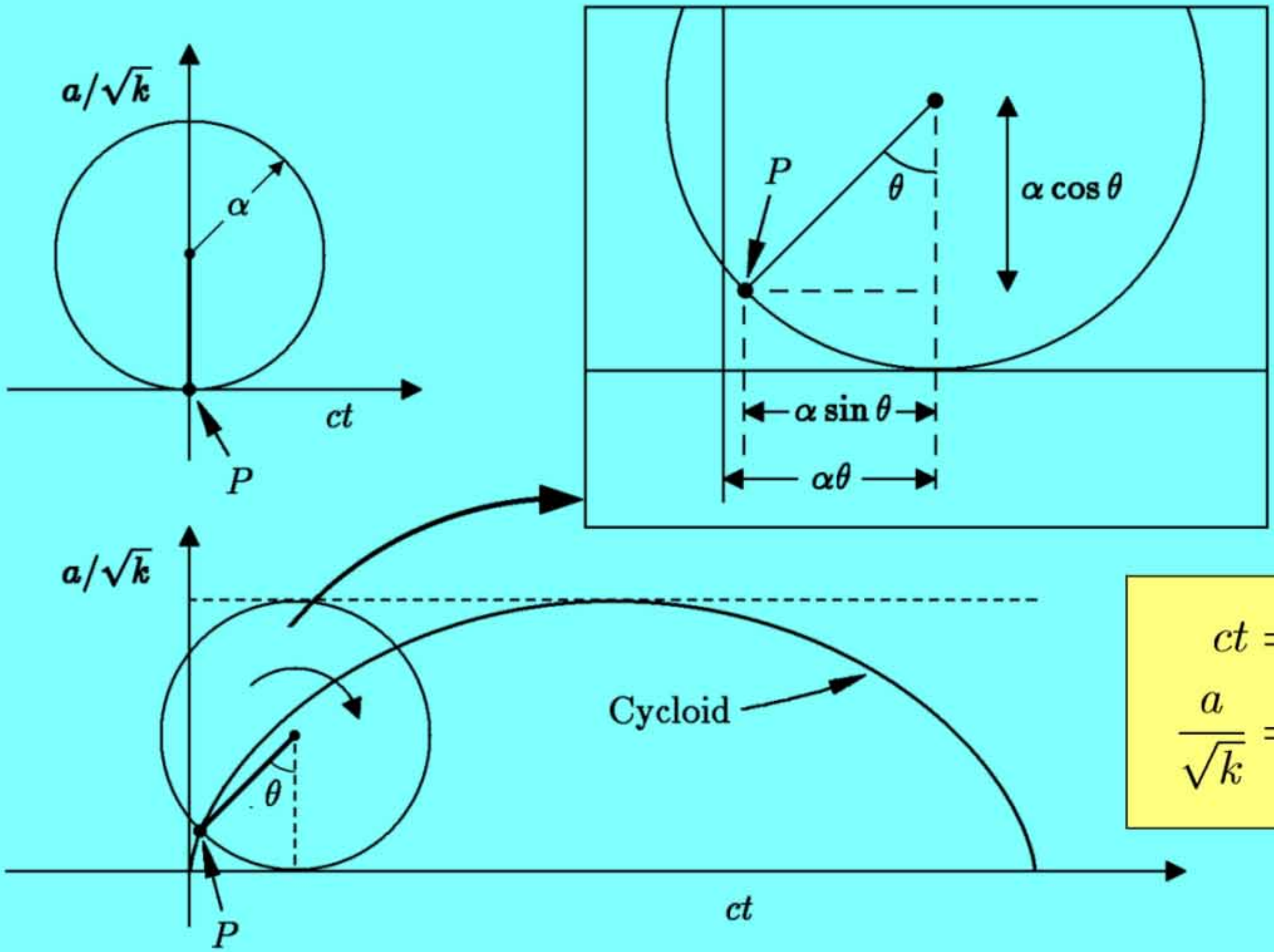
$$\tilde{a}_f = \alpha(1-\cos\theta_f) \ .$$

Parametric Solution for the Evolution of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta) \ ,$$

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

The angle θ is sometimes called the "development angle," because it describes the stage of development of the universe. The universe begins at $\theta = 0$, reaches its maximum expansion at $\theta = \pi$, and then is terminated by a big crunch at $\theta = 2\pi$.



$$ct = \alpha(\theta - \sin\theta) \ ,$$
$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

Duration and Maximum Size

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \qquad \Longrightarrow \qquad \boxed{\frac{a_{\max}}{\sqrt{k}} = 2\alpha\ ,}$$

where

$$\alpha = \frac{4\pi}{3} \frac{G\rho a^3}{k^{3/2} c^2}\ .$$

Similarly, $ct = \alpha(\theta - \sin\theta)$ implies that the total duration of the universe, from big bang to big crunch is

$$\boxed{t_{\text{total}} = \frac{2\pi\alpha}{c} = \frac{\pi a_{\max}}{c\sqrt{k}}\ .}$$

Age of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta)$$

gives the age in terms of α and θ. But astronomers measure H and Ω. So we would like to express the age in terms of H and Ω.

Start with ρ:

$$\rho = \Omega\rho_c = \left(\frac{3H^2}{8\pi G}\right)\Omega\ .$$

The first-order Friedmann equation can then be rewritten as

$$H^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} \quad \Longrightarrow \quad H^2 = H^2\Omega - \frac{kc^2}{a^2}\ ,$$

so

$$\boxed{\tilde{a} = \frac{a}{\sqrt{k}} = \frac{c}{|H|\sqrt{\Omega - 1}}\ .}$$

8.286 Lecture 8
October 3, 2022

THE DYNAMICS OF NEWTONIAN COSMOLOGY, PART 4

INTRODUCTION TO NON-EUCLIDEAN SPACES, PART 1

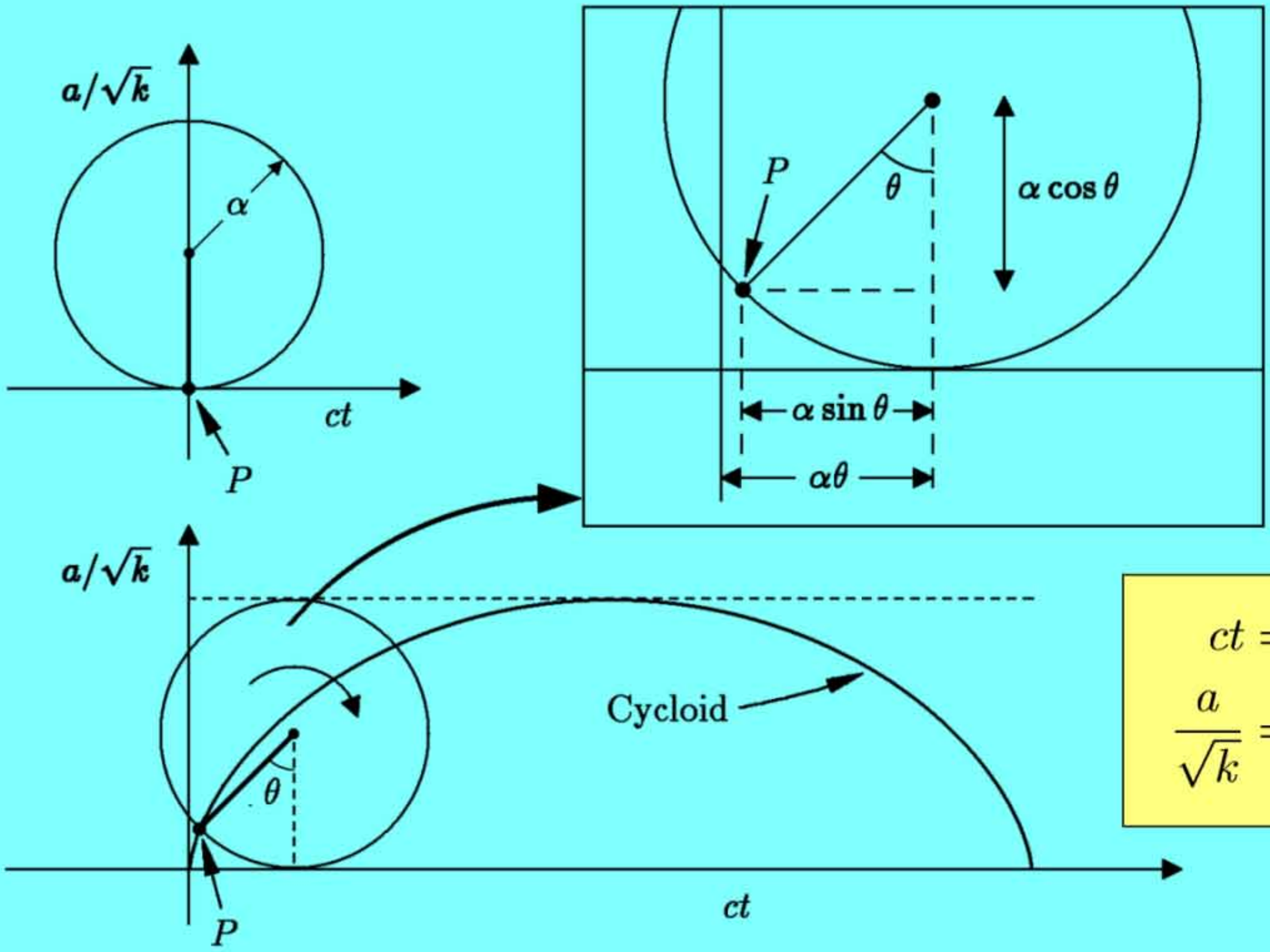
Parametric Solution for the Evolution of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta) \ ,$$

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

The angle θ is sometimes called the "development angle," because it describes the stage of development of the universe. The universe begins at $\theta = 0$, reaches its maximum expansion at $\theta = \pi$, and then is terminated by a big crunch at $\theta = 2\pi$.

Review from the previous lecture



$$ct = \alpha(\theta - \sin\theta) \ ,$$

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

Duration and Maximum Size

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \quad \Longrightarrow \quad \boxed{\frac{a_{\rm max}}{\sqrt{k}} = 2\alpha \ ,}$$

where

$$\alpha = \frac{4\pi}{3} \frac{G\rho a^3}{k^{3/2} c^2} \ .$$

Similarly, $ct = \alpha(\theta - \sin\theta)$ implies that the total duration of the universe, from big bang to big crunch is

$$\boxed{t_{\rm total} = \frac{2\pi\alpha}{c} = \frac{\pi a_{\rm max}}{c\sqrt{k}} \ .}$$

Age of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta)$$

gives the age in terms of α and θ. But astronomers measure H and Ω. So we would like to express the age in terms of H and Ω.

Start with ρ:

$$\rho = \Omega\rho_c = \left(\frac{3H^2}{8\pi G}\right)\Omega\ .$$

The first-order Friedmann equation can then be rewritten as

$$H^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} \quad \Longrightarrow \quad H^2 = H^2\Omega - \frac{kc^2}{a^2}\ ,$$

so

$$\boxed{\tilde{a} = \frac{a}{\sqrt{k}} = \frac{c}{|H|\sqrt{\Omega - 1}}\ .}$$

$$\boxed{\tilde{a} = \frac{a}{\sqrt{k}} = \frac{c}{|H|\sqrt{\Omega - 1}}\ .}$$

In taking the square root, recall that $a > 0$, $k > 0$, while H changes sign — it is positive during the expansion phase, and negative during the collapse phase. So we need $|H|$, not just H, for the equation to be valid. Then

$$\alpha = \frac{4\pi}{3} \frac{G\rho\tilde{a}^3}{c^2} = \frac{c}{2|H|} \frac{\Omega}{(\Omega - 1)^{3/2}}\ .$$

To find the age, we need to express α and θ in terms of H and Ω. To express θ, use expression for $\tilde{a}$ above, and 2nd parametric equation

$$\tilde{a} = \frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta)\ .$$

Then

$$\frac{c}{|H|\sqrt{\Omega - 1}} = \frac{c}{2|H|} \frac{\Omega}{(\Omega - 1)^{3/2}}(1 - \cos\theta)\ ,$$

Then

$$\frac{c}{|H|\sqrt{\Omega - 1}} = \frac{c}{2|H|}\,\frac{\Omega}{(\Omega - 1)^{3/2}}(1 - \cos\theta) \ ,$$

which can be solved for either $\cos\theta$ or for Ω:

$$\boxed{\cos\theta = \frac{2 - \Omega}{\Omega} \ , \quad \Omega = \frac{2}{1 + \cos\theta} \ .}$$

Evolution of Ω: At $t = 0$, $\theta = 0$, so $\Omega = 1$. Any (matter-dominated) closed universe begins with $\Omega = 1$.

As θ increases from 0 to π, Ω grows from 1 to infinity. At $\theta = \pi$, a reaches its maximum size, and $H = 0$. So $\rho_c = 0$ and $\Omega = \infty$.

During the collapse phase, $\pi < \theta < 2\pi$, Ω falls from ∞ to 1.

What about $\sin\theta$?

$$\sin\theta = \pm\sqrt{1-\cos^2\theta} = \pm\frac{2\sqrt{\Omega-1}}{\Omega}\ .$$

$\sin\theta$ is positive during the expansion phase (while $0 < \theta < \pi$), and negative during the collapse phase (while $\pi < \theta < 2\pi$).

Evolution of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta) \ ,$$

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

Evolution of a Closed Matter-Dominated Universe

$$ct = \alpha(\theta - \sin\theta) \ ,$$

$$\frac{a}{\sqrt{k}} = \alpha(1 - \cos\theta) \ .$$

$$t = \frac{\Omega}{2|H|(\Omega - 1)^{3/2}} \left\{ \arcsin\left(\pm \frac{2\sqrt{\Omega - 1}}{\Omega} \right) \mp \frac{2\sqrt{\Omega - 1}}{\Omega} \right\} \ .$$

$$t = \frac{\Omega}{2|H|(\Omega-1)^{3/2}} \left\{ \sin^{-1}\left(\pm \frac{2\sqrt{\Omega-1}}{\Omega} \right) \mp \frac{2\sqrt{\Omega-1}}{\Omega} \right\} .$$

Quadrant	$\theta = \sin^{-1}()$	Phase	Ω	Sign Choice
1	0 to $\frac{\pi}{2}$	Expanding	1 to 2	Upper
2	$\frac{\pi}{2}$ to π	Expanding	2 to ∞	Upper
3	π to $\frac{3\pi}{2}$	Contracting	∞ to 2	Lower
4	$\frac{3\pi}{2}$ to 2π	Contracting	2 to 1	Lower

Evolution of Open Matter-Dominated Universes

$$ct = \alpha(\sinh\theta - \theta) \; ,$$

$$\frac{a}{\sqrt{\kappa}} = \alpha(\cosh\theta - 1) \; .$$

where $\kappa = -k$, and

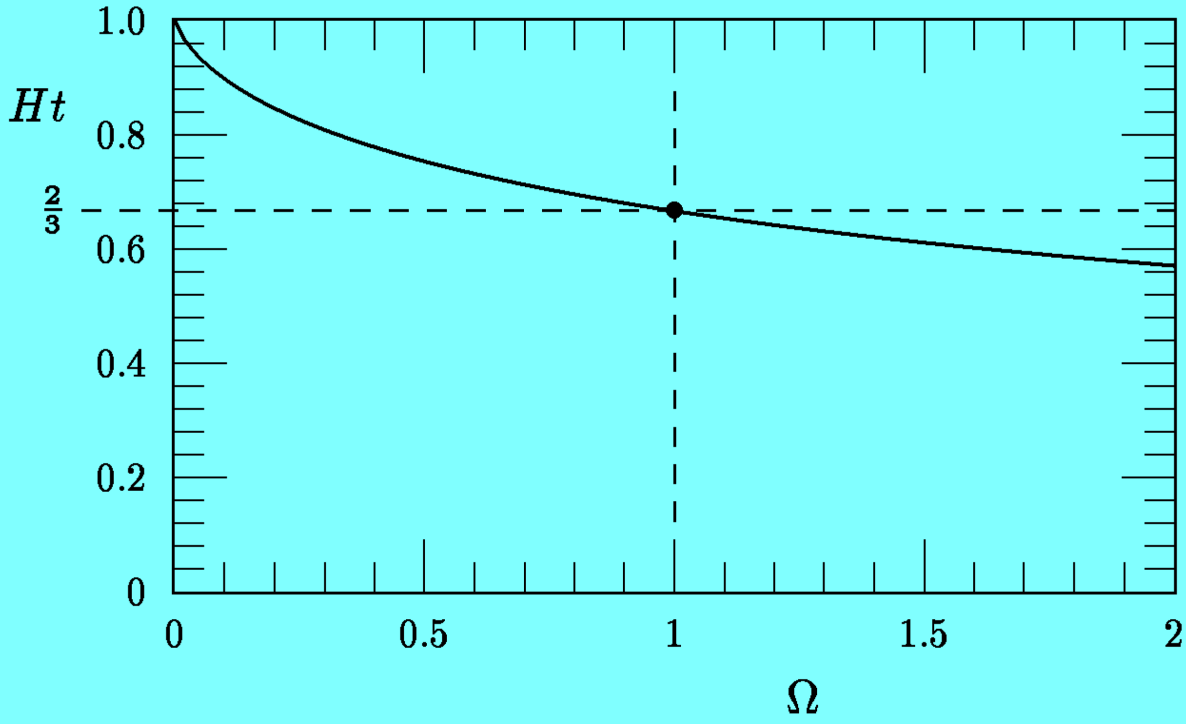
$$\tilde{a}(t) = \frac{a(t)}{\sqrt{\kappa}} \; , \qquad \alpha \equiv \frac{4\pi}{3}\frac{G\rho\tilde{a}^3}{c^2} \; .$$

θ evolves from 0 to ∞.

Age for Open, Flat, and Closed Matter-Dominated Universes

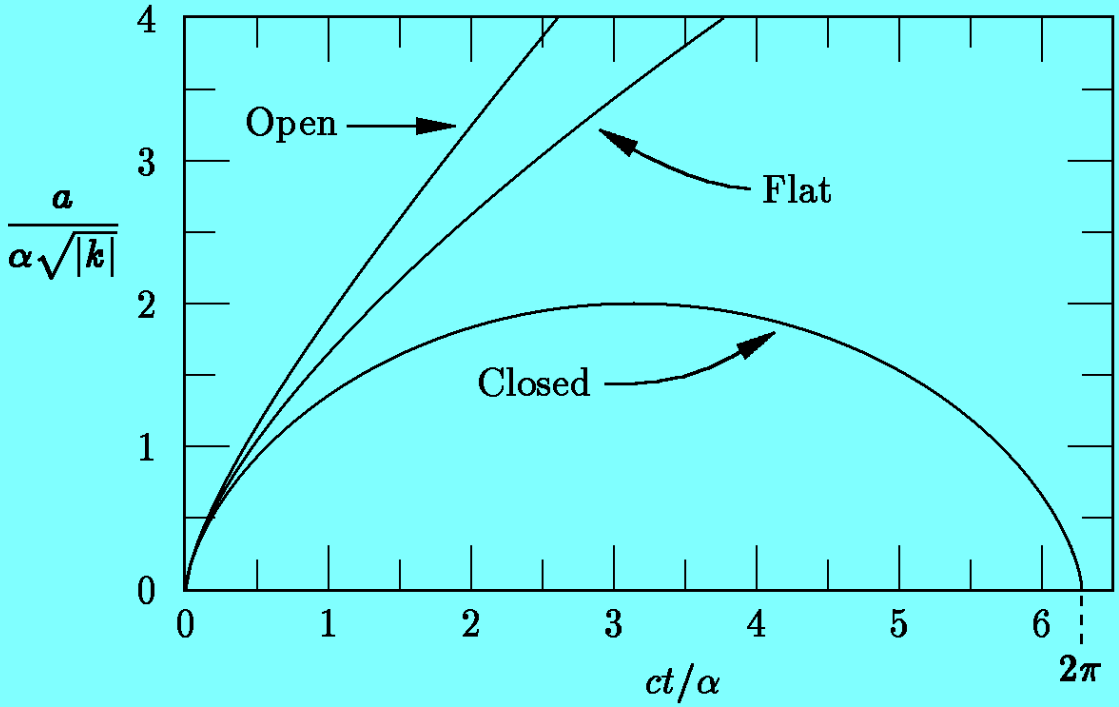
$$|H|\,t = \begin{cases} \dfrac{\Omega}{2(1-\Omega)^{3/2}}\left[\dfrac{2\sqrt{1-\Omega}}{\Omega} - \sinh^{-1}\left(\dfrac{2\sqrt{1-\Omega}}{\Omega}\right)\right] & \text{if } \Omega < 1 \\ 2/3 & \text{if } \Omega = 1 \\ \dfrac{\Omega}{2(\Omega-1)^{3/2}}\left[\sin^{-1}\left(\pm\dfrac{2\sqrt{\Omega-1}}{\Omega}\right) \mp \dfrac{2\sqrt{\Omega-1}}{\Omega}\right] & \text{if } \Omega > 1 \end{cases}$$

The Age of a Matter-Dominated Universe



The age of a matter-dominated universe, expressed as Ht (where t is the age and H is the Hubble expansion rate), as a function of Ω. The curve describes all three cases of an open ($\Omega < 1$), flat ($\Omega = 1$), and closed ($\Omega > 1$) universe.

Evolution of a Matter-Dominated Universe



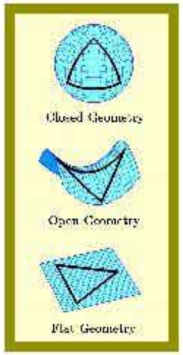
The evolution of a matter-dominated universe. Closed and open universes can be characterized by a single parameter α. With the scalings shown on the axis labels, the evolution of a matter-dominated universe is described in all cases by the curves shown in this graph.

8.286 Lecture 8, Part 2
October 3, 2022

INTRODUCTION TO NON-EUCLIDEAN SPACES

Ants on a Pringle

Mustafa Amin
18.10.2011

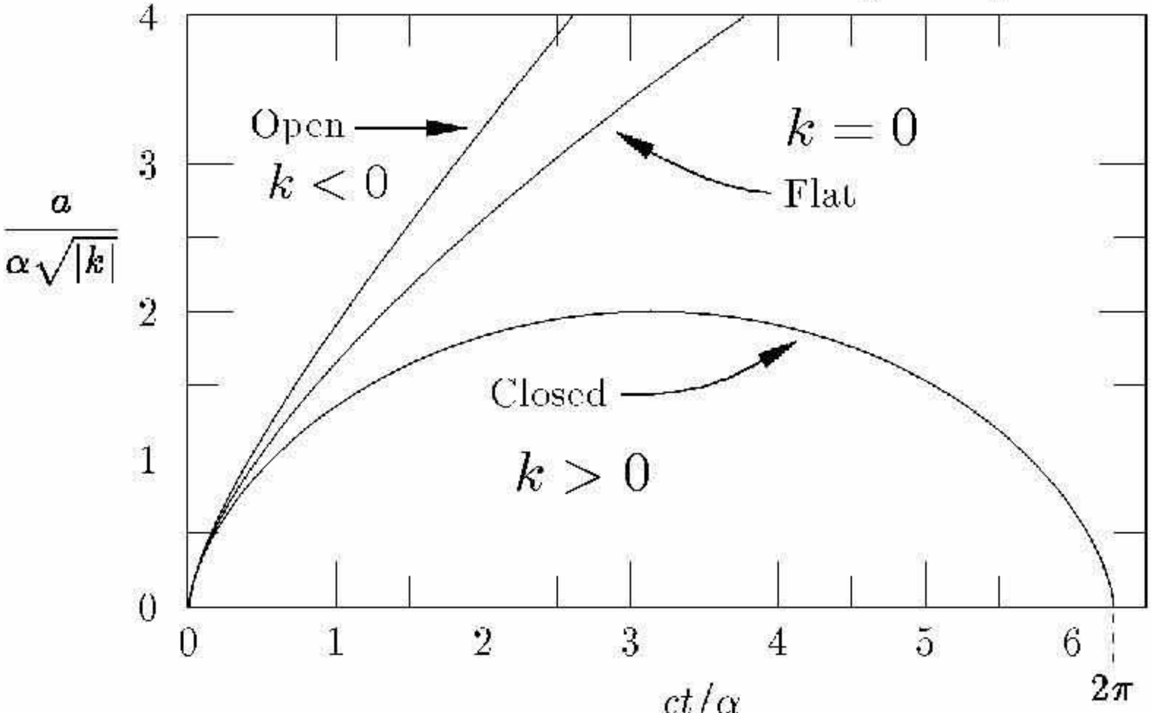


Slide created by Mustafa Amin

Recap

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}$$

Image: courtesy Alan Guth



Note: matter dominated universes only!

k ?

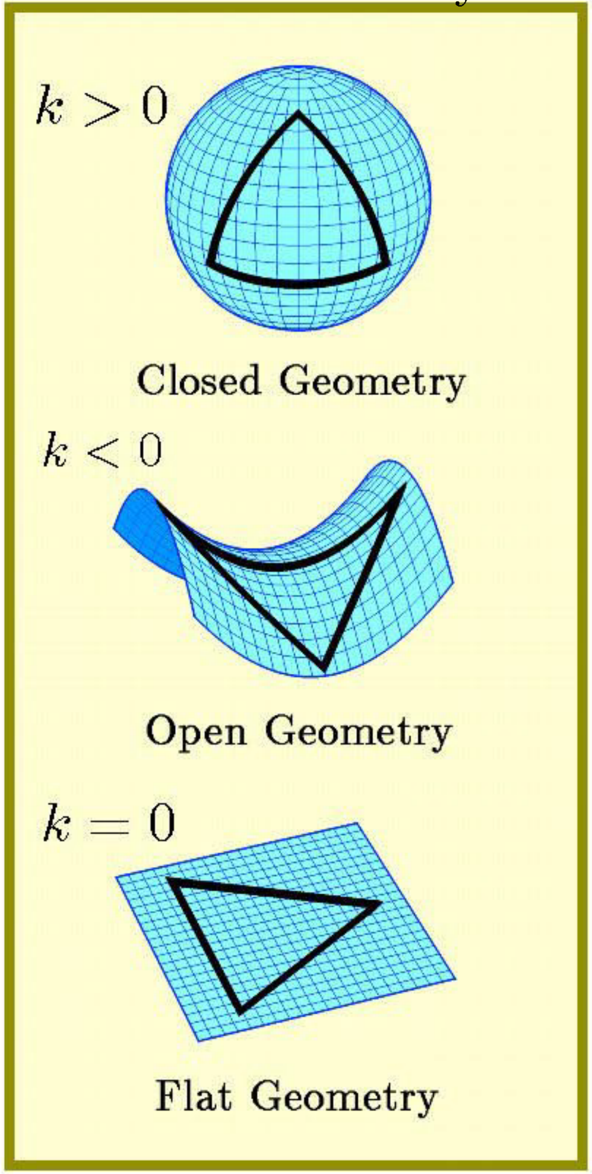
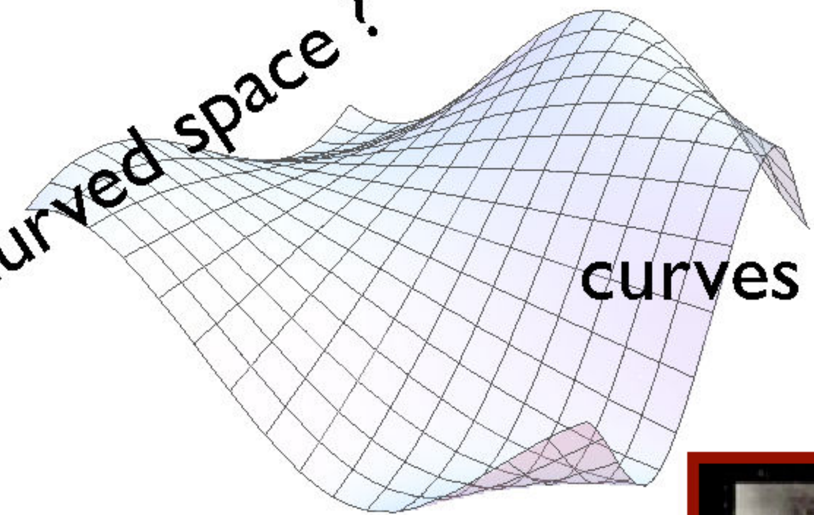


Image: courtesy Alan Guth

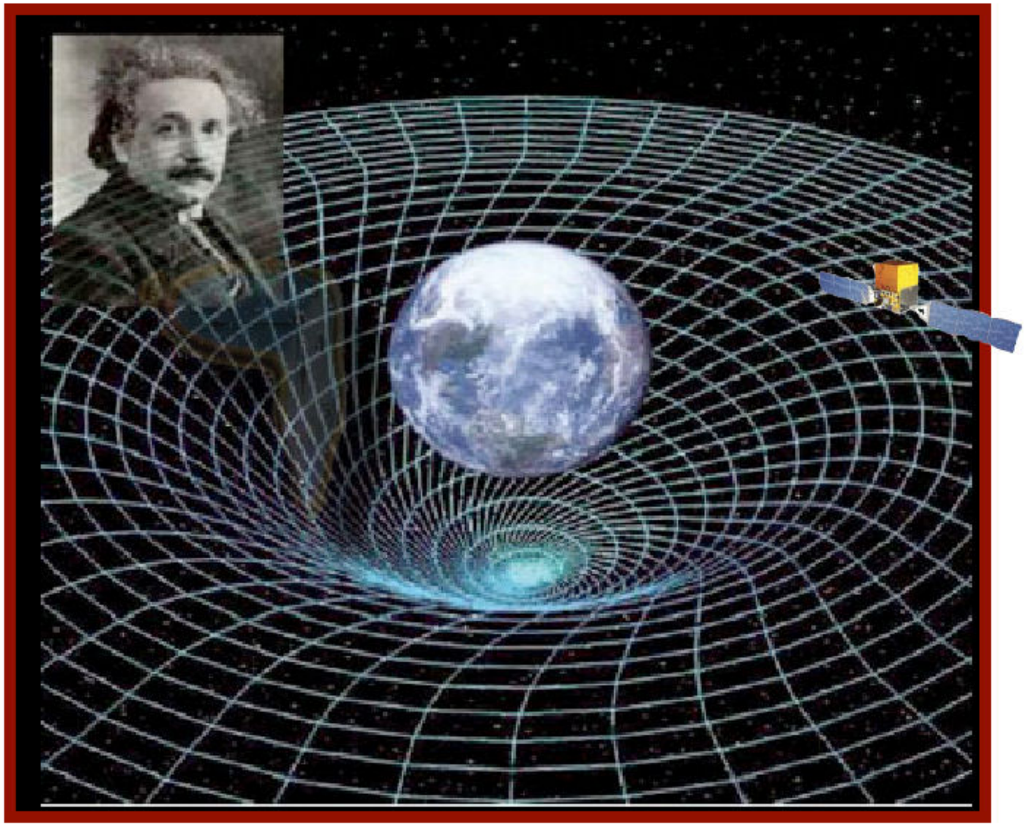
Slide created by Mustafa Amin



curved space ?

curves with respect to what ?

curved spacetime ?



non-Euclidean geometry

Note on image: I have added the "Dali" clock and Fermi satellite images to the original image created for the Gravity Probe B collaboration

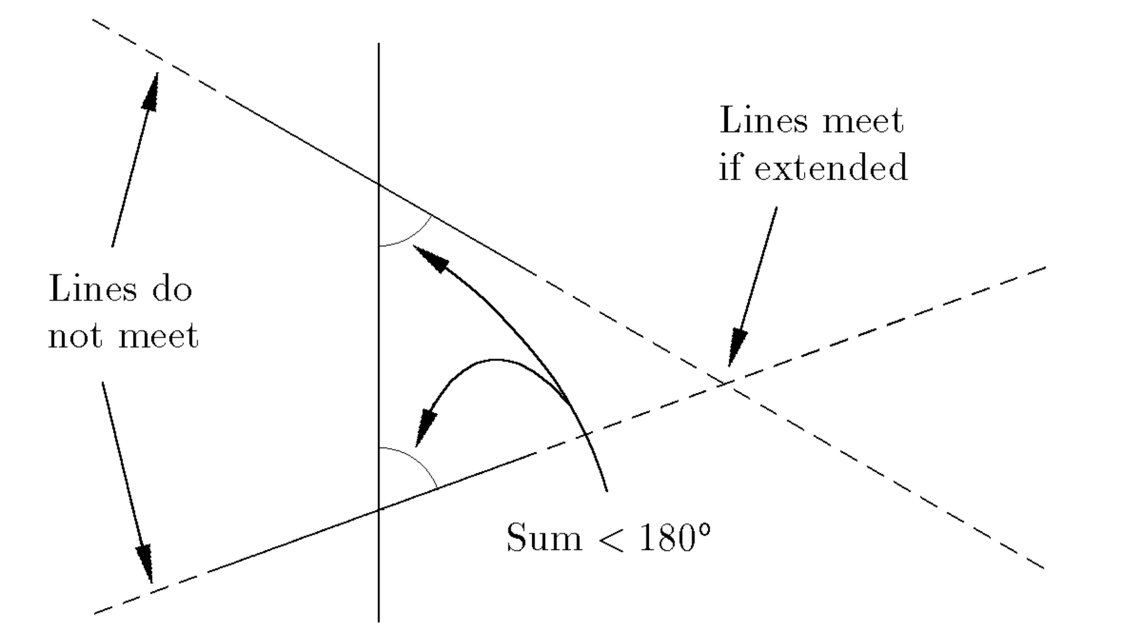
Euclid's *Postulates*

1. A straight line segment can be drawn joining any two points.
2. Any straight line segment can be extended indefinitely in a straight line.
3. Given a straight line segment , a circle can be drawn having the segment as radius and one endpoint as center.
4. All right angles are congruent.

5. If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines if produced indefinitely meet on that side on which the angles are less than two right angles

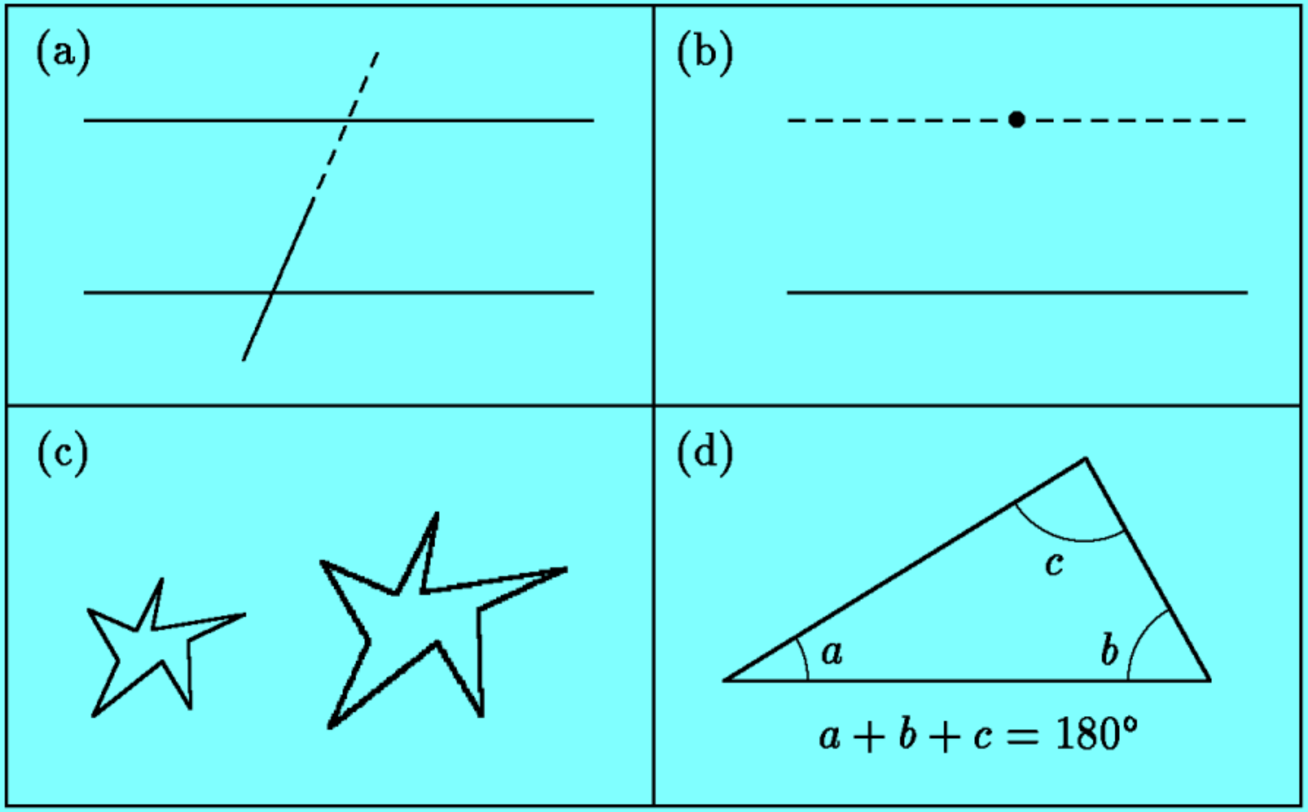
Corrected 10/10/13

Euclid's Fifth Postulate



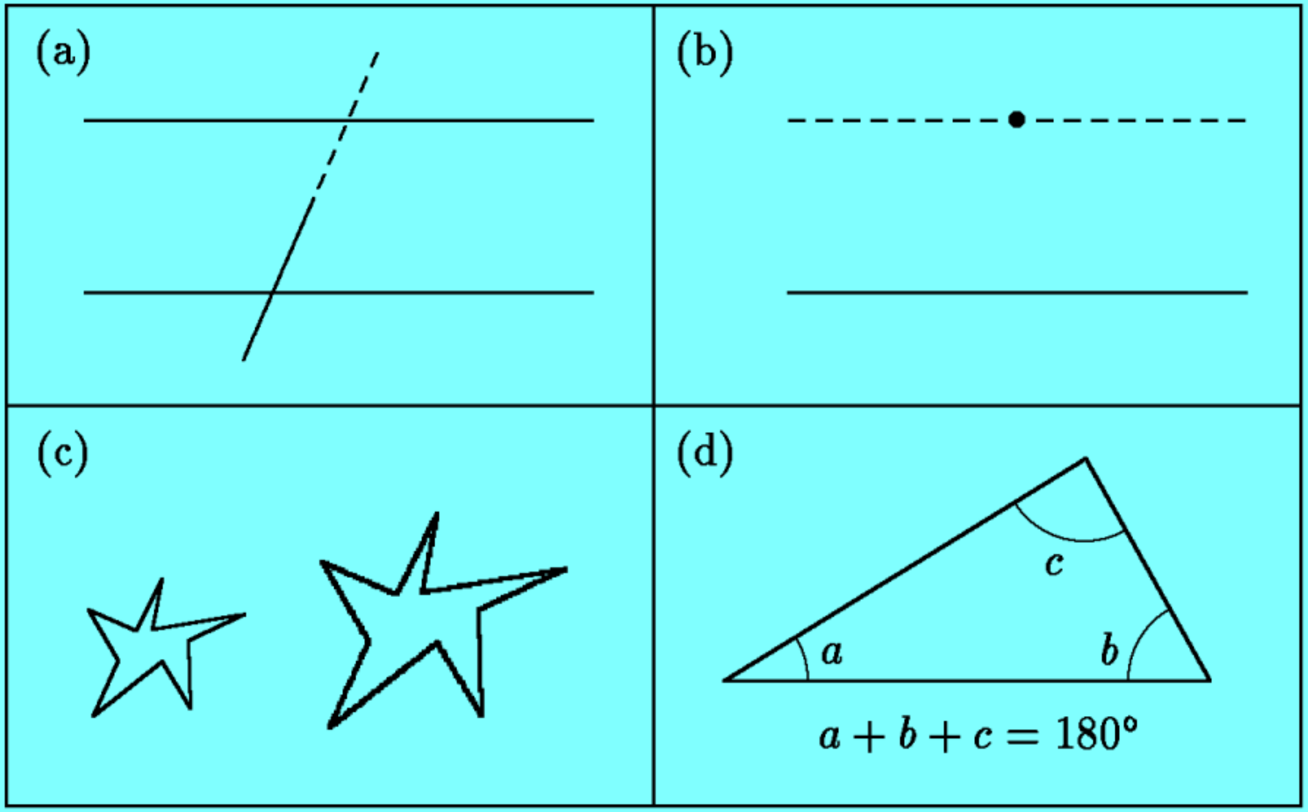
"If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines if produced indefinitely meet on that side on which the angles are less than two right angles." *[This statement is interpreted to imply that the two straight lines will never meet if extended on the opposite side.]*

Equivalent Statements of the 5th Postulate



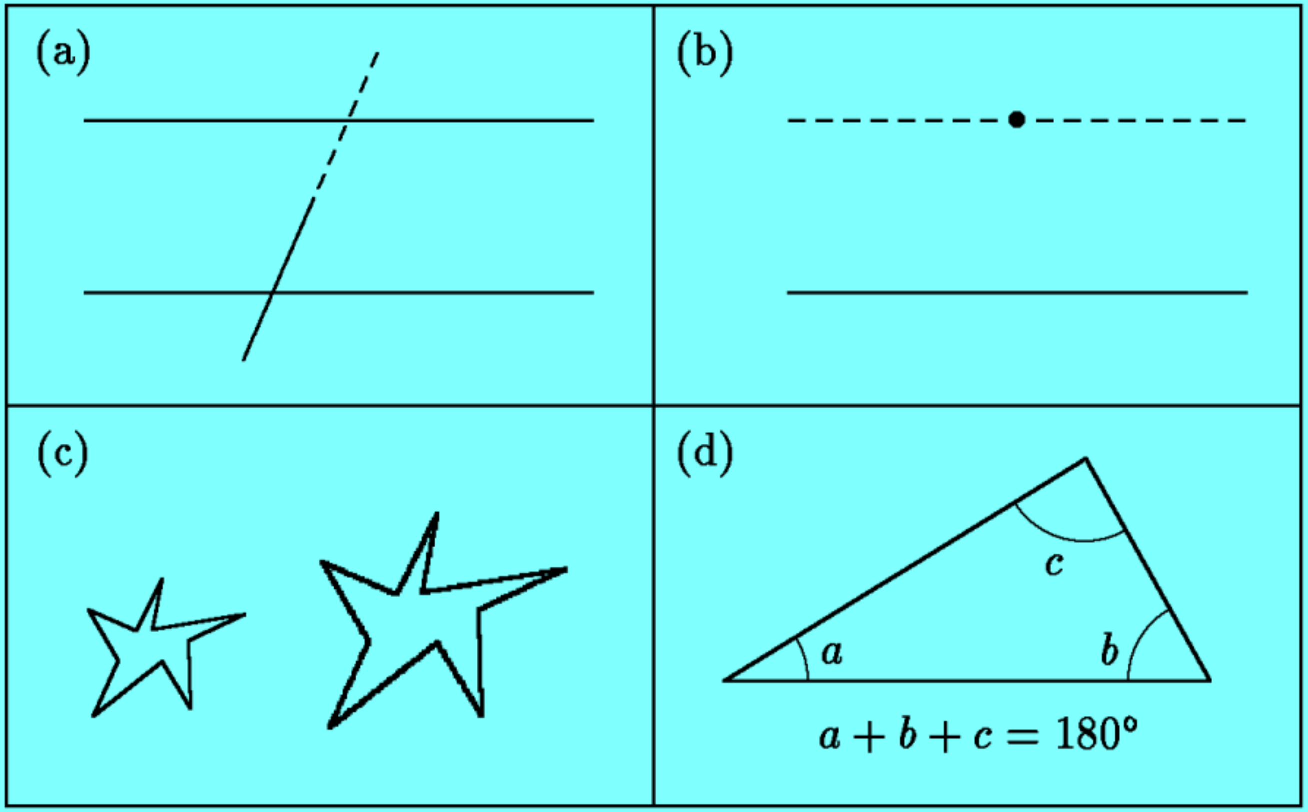
(a) "If a straight line intersects one of two parallels (i.e, lines which do not intersect however far they are extended), it will intersect the other also."

Equivalent Statements of the 5th Postulate



(b) "There is one and only one line that passes through any given point and is parallel to a given line."

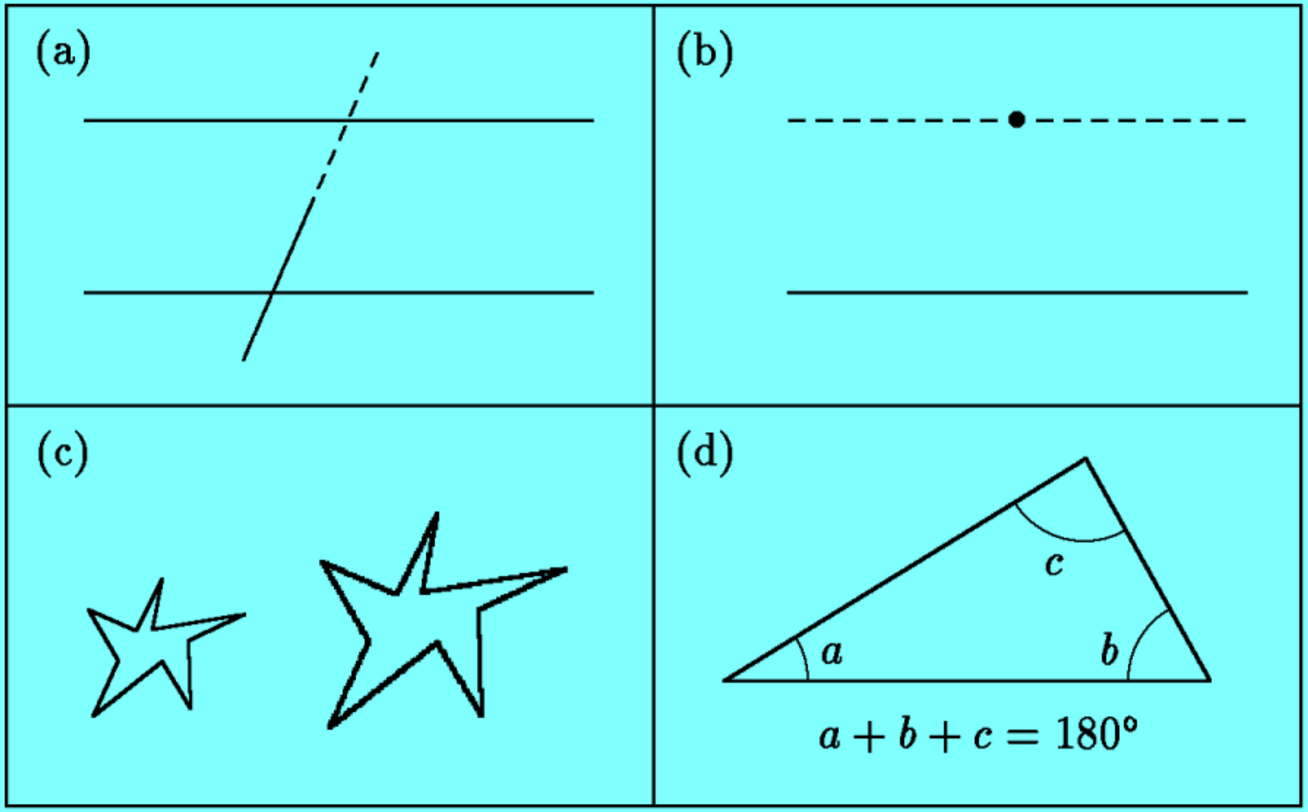
Equivalent Statements of the 5th Postulate



(c) "Given any figure there exists a figure, similar to it, of any size." (Two polygons are similar if their corresponding angles are equal, and their corresponding sides are proportional.)

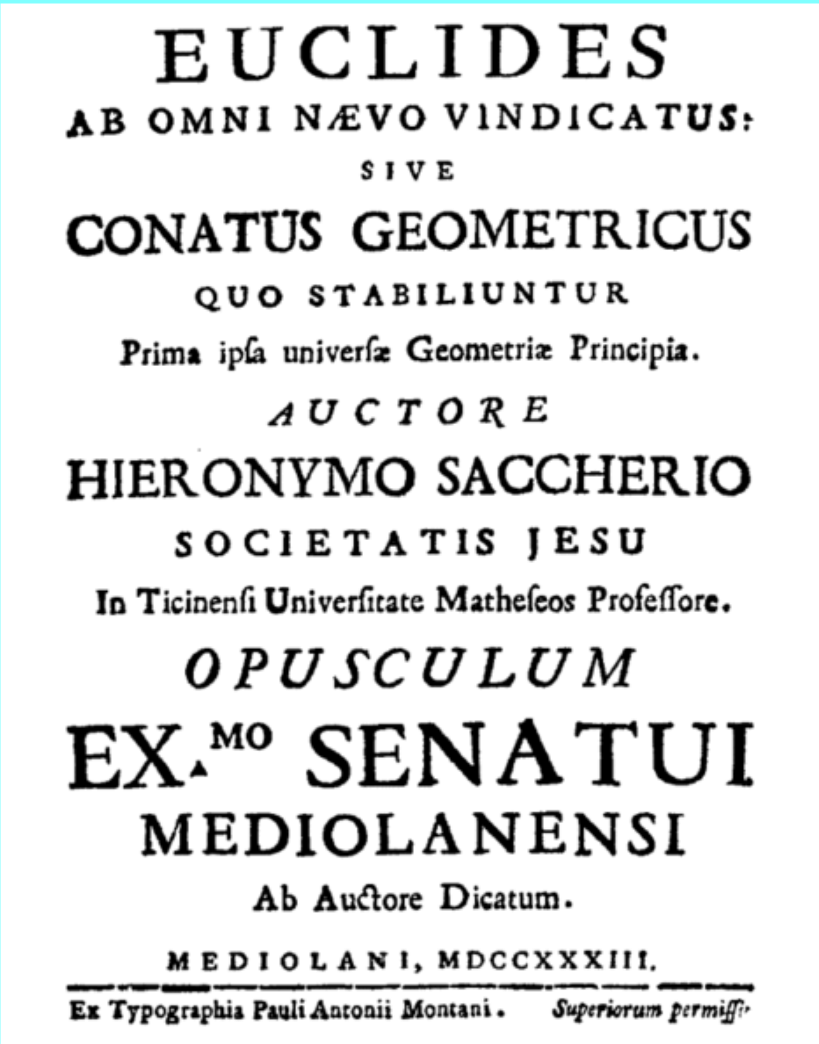
Alan Guth
Massachusetts Institute of Technology

Equivalent Statements of the 5th Postulate



(d) "There is a triangle in which the sum of the three angles is equal to two right angles (i.e., 180°)."

Giovanni Geralamo Saccheri (1667–1733)



EUCLIDES
AB OMNI NÆVO VINDICATUS:
SIVE
CONATUS GEOMETRICUS
QUO STABILIUNTUR
Prima ipsa universæ Geometriæ Principia.
AUCTORE
HIERONYMO SACCHERIO
SOCIETATIS JESU
In Ticinensi Universitate Matheseos Professore.
OPUSCULUM
EX.MO SENATUI
MEDIOLANENSI
Ab Auctore Dicatum.
MEDIOLANI, MDCCXXXIII.
Ex Typographia Pauli Antonii Montani. *Superiorum permissu.*

In 1733, Saccheri, a Jesuit priest, published *Euclides ab omni naevo vindicatus (Euclid Freed of Every Flaw).*

The book was a study of what geometry would be like if the 5th postulate were false.

He hoped to find an inconsistency, but failed.

Carl Friedrich Gauss (1777–1855)

German mathematician and physicist.

Born as the son of a poor working-class parents. His mother was illiterate and never even recorded the date of his birth.

His students included Richard Dedekind, Bernhard Riemann, Peter Gustav Lejeune Dirichlet, Gustav Kirchhoff, and August Ferinand Möbius.

János Bolyai (1802–1860)

Hungarian mathematician and army officer.

Son of Farkas Bolyai, a teacher of mathematics, physics, and chemistry at the Calvinist College in Marosvásárhely, Hungary.

Attended Marosvásárhely College and later studied military engineering at the Academy of Engineering at Vienna, because that is what his family could afford.

Served 11 years in the army engineering corps; during this time he developed his non-Euclidean geometry, which was published as an appendix to a book written by his father.

Retired from the army at age 31 due to poor health, and died in relative poverty at age 57, from pneumonia.

Nikolai Ivanovich Lobachevsky (1792–1856)

Russian mathematician and college teacher.

Born in Russia from Polish parents; father was a clerk in a land-surveying office, but died when Nikolai was only seven.

Moved to Kazan, attending Kazan Gymnasium and later was given a scholarship to Kazan University. He remained at Kazan University on the faculty.

Work on non-Euclidean geometry published in the *Kazan Messenger* in 1829, but was rejected for publication by the St. Petersburg Academy of Sciences.

He was "asked to retire" at age 54, and died 10 years later in poor health and in poverty. His work was never appreciated during his lifetime.

~1750-1850

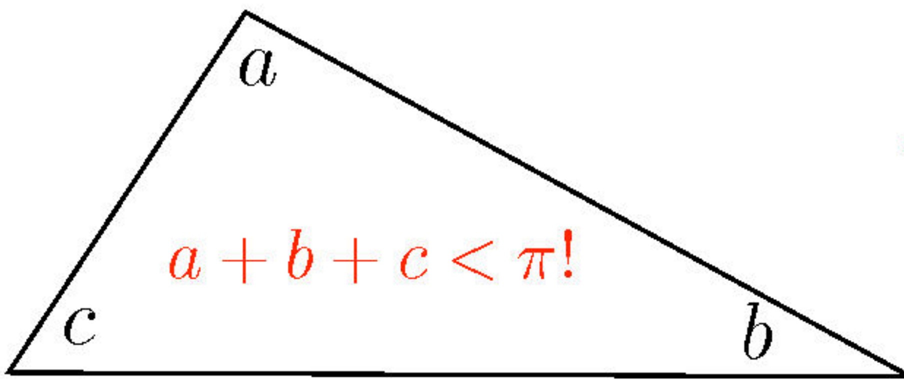
Bolyai

Lobachevski

Gauss

- infinite
- constant negative curvature

~~5th Postulate~~



Slide created by Mustafa Amin

GBL geometry with Klein

1. constant negative curvature
2. infinite
3. ~~5th postulate~~

(x_1, y_1)

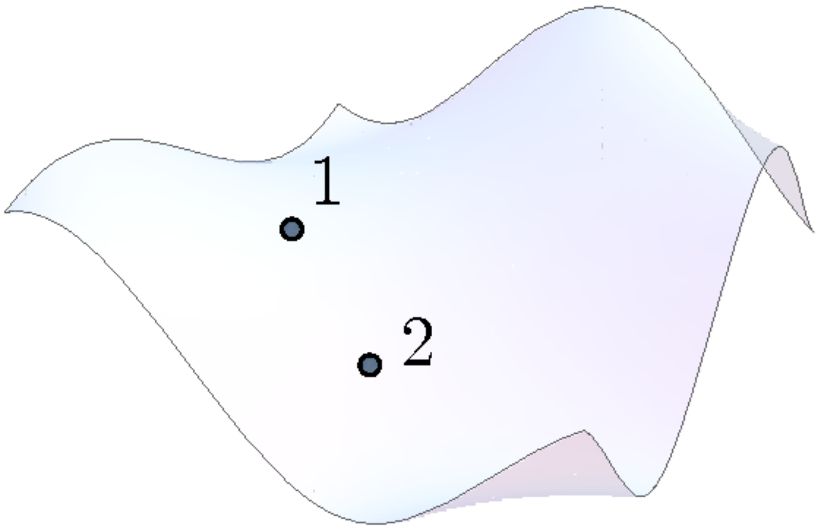
(x_2, y_2)

$$x^2 + y^2 < 1$$

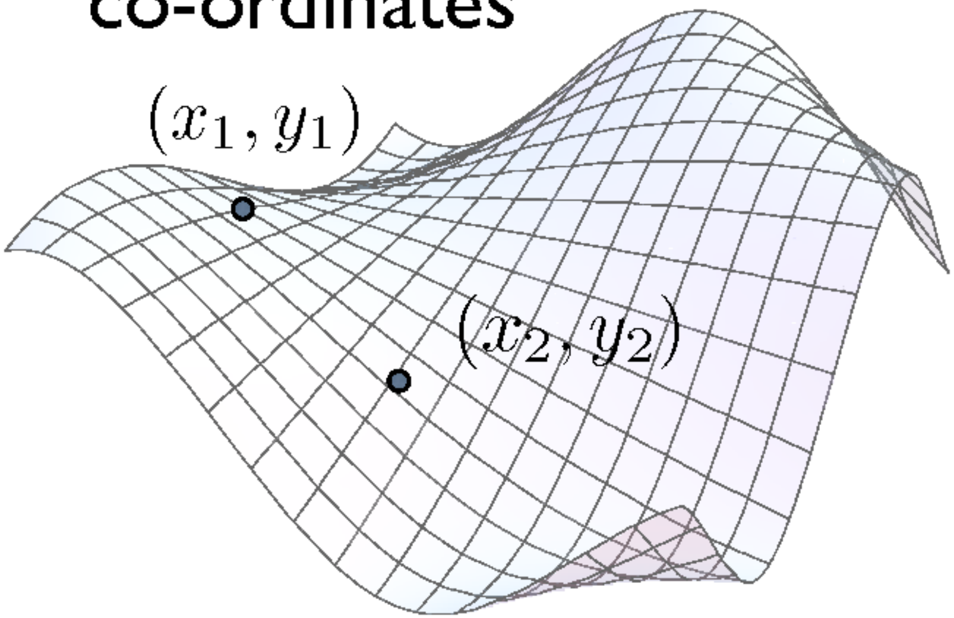
$$d(1,2) = a \cosh^{-1}\left[\frac{1 - x_1 x_2 - y_1 y_2}{\sqrt{1 - x_1^2 - y_1^2}\sqrt{1 - x_2^2 - y_2^2}}\right]$$

Note: no global embedding in 3D Euclidean space possible

Geometry (after Klein)



co-ordinates



distance function

$$d[(x_1, y_1), (x_2, y_2)]$$

Intrinsic Geometry

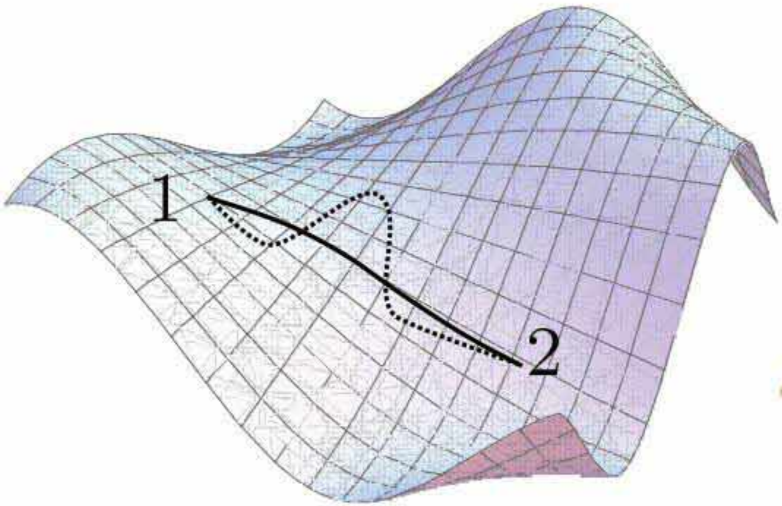


Image of ant from: http://www.termiteterry.com/

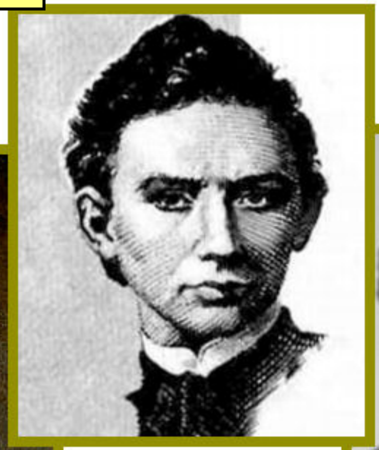
8.286 Class 10
October 12, 2022

INTRODUCTION TO NON-EUCLIDEAN SPACES, PART 2

Review from the previous lecture

~1750-1850

Gauss

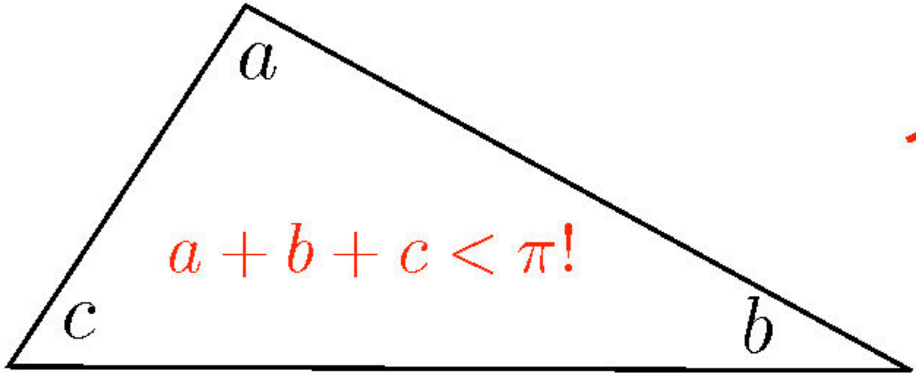


Bolyai

Lobachevski

- infinite
- constant negative curvature

~~5th Postulate~~



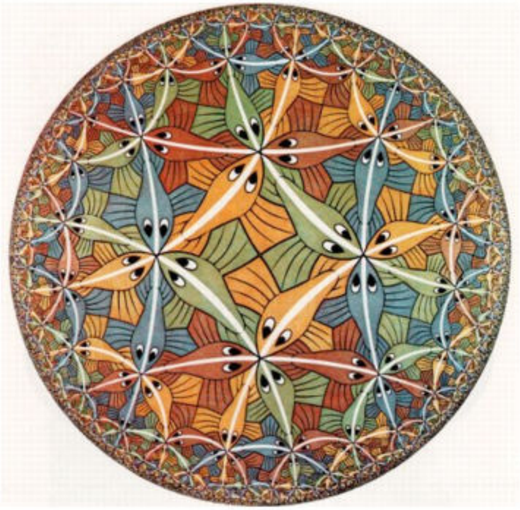
Slide created by Mustafa Amin

Review from the previous lecture

Slide created by Mustafa Amin

GBL geometry with Klein

1. constant negative curvature
2. infinite
3. ~~5th postulate~~



(x_1, y_1)

(x_2, y_2)

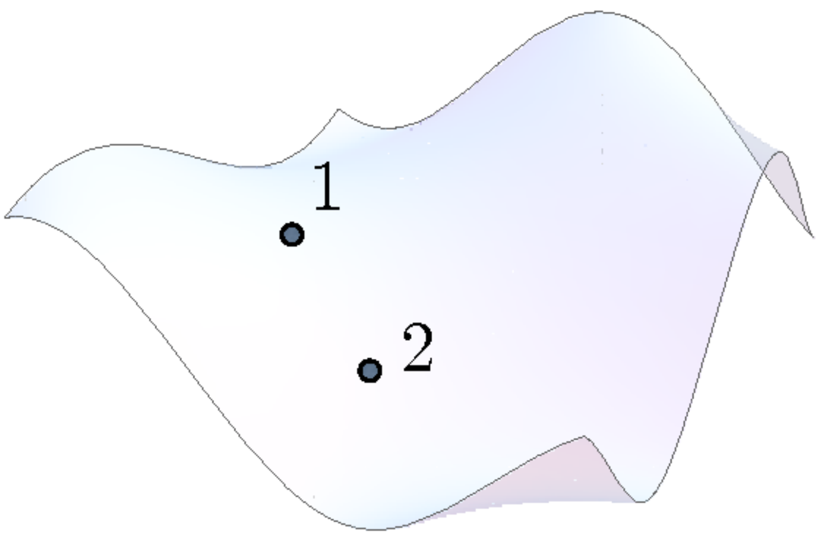
$$x^2 + y^2 < 1$$

$$d(1,2) = a \cosh^{-1}\left[\frac{1 - x_1 x_2 - y_1 y_2}{\sqrt{1 - x_1^2 - y_1^2}\sqrt{1 - x_2^2 - y_2^2}}\right]$$

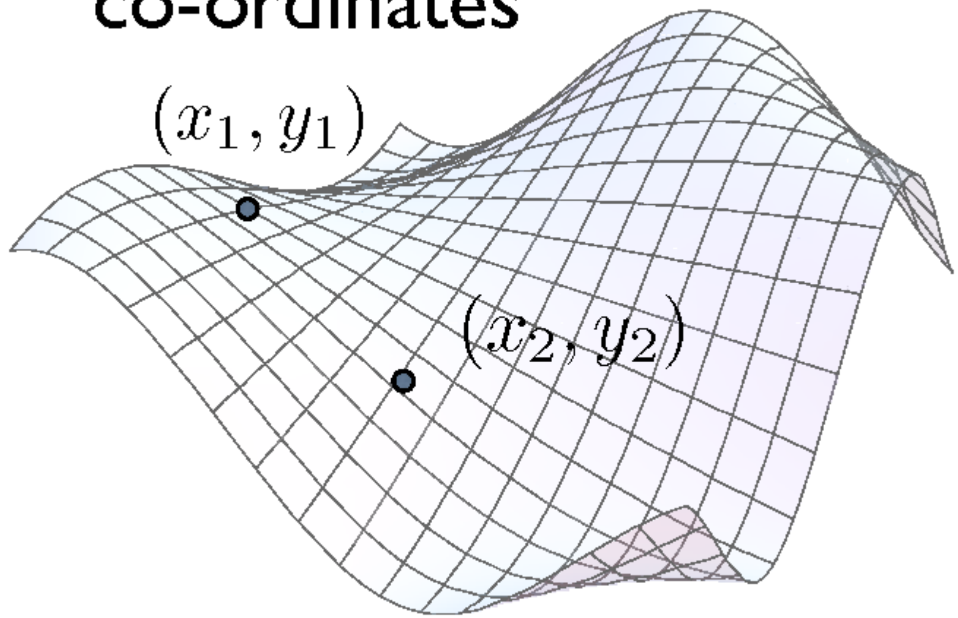
Note: no global embedding in 3D Euclidean space possible

Slide created by Mustafa Amin

Geometry (after Klein)



co-ordinates



distance function

$$d[(x_1, y_1), (x_2, y_2)]$$

Review from the previous lecture

Slide created by Mustafa Amin

Intrinsic Geometry

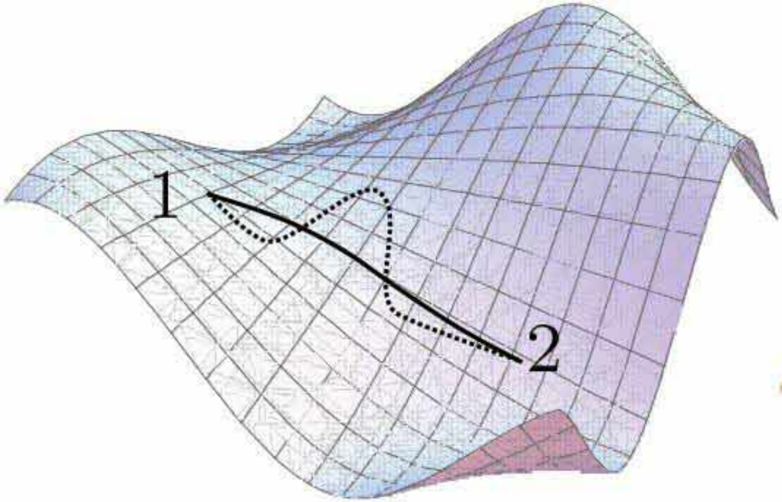
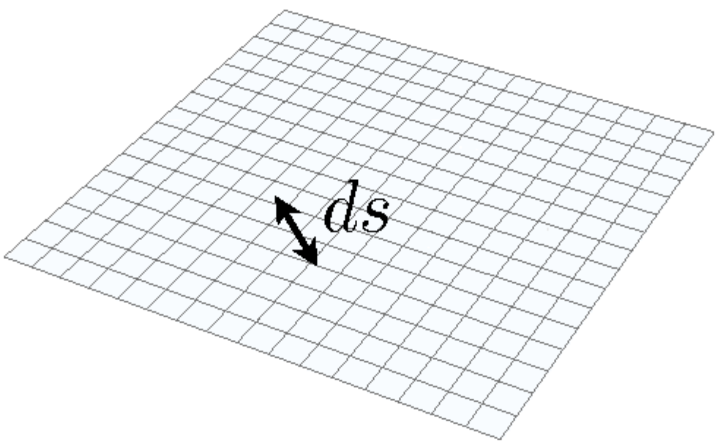
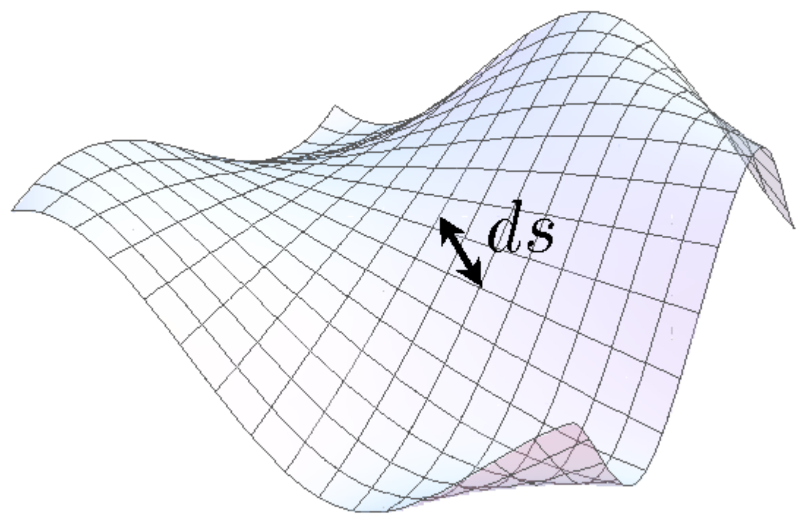


Image of ant from: http://www.termiteterry.com/

Slide created by Mustafa Amin

tiny distances

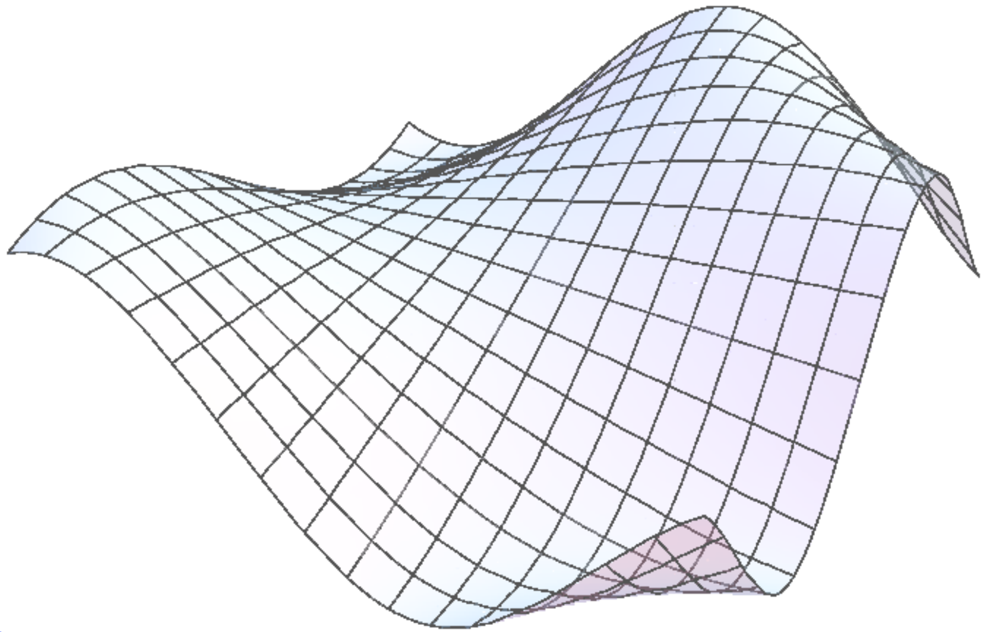




$$ds^2 = dx^2 + dy^2 \qquad ds^2 = g_{xx}(x,y)dx^2 + 2g_{xy}(x,y)dxdy + g_{yy}(x,y)dy^2$$

quadratic form

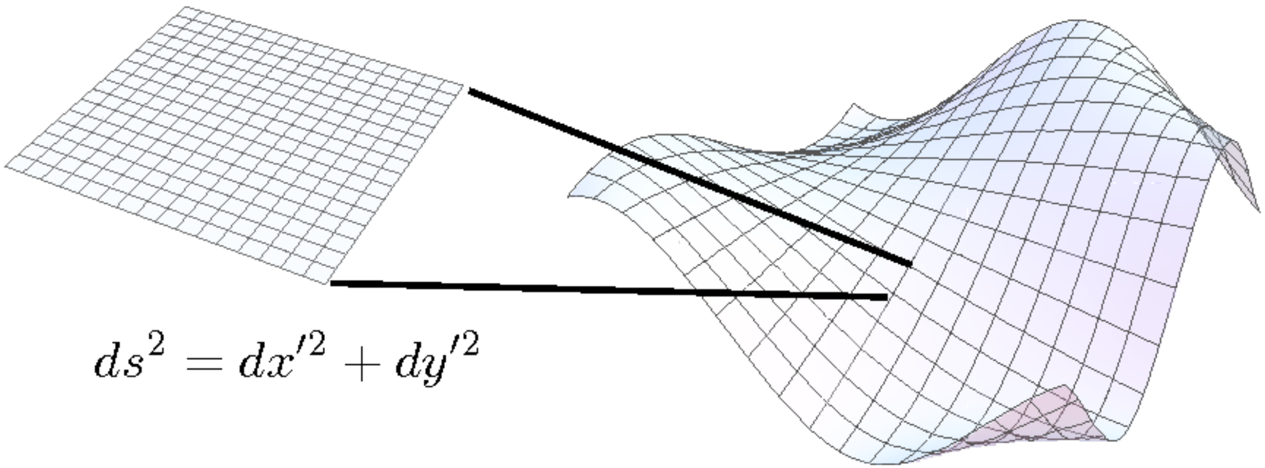
Image:www.easternct.edu/career/webresources.htm



$$ds^2 = g_{xx}(x,y)dx^2 + 2g_{xy}(x,y)dxdy + g_{yy}(x,y)dy^2$$

Slide created by Mustafa Amin

locally Euclidean

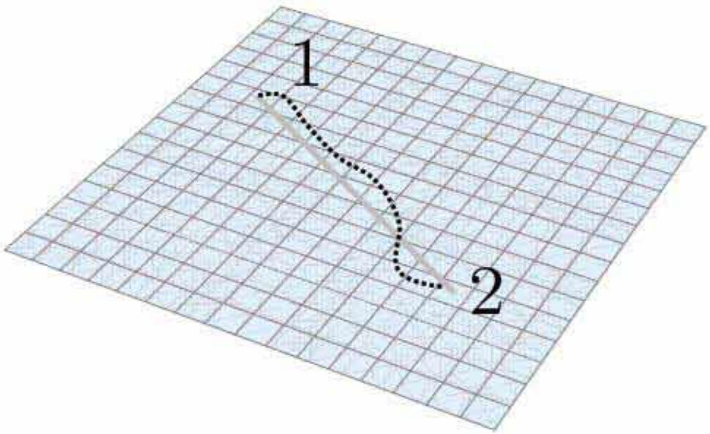


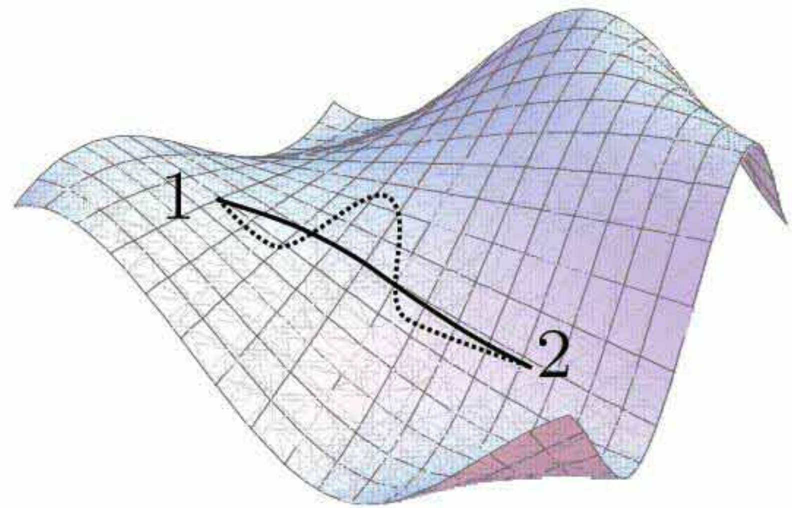
$$ds^2 = dx'^2 + dy'^2$$

$$ds^2 = g_{xx}(x, y)dx^2 + 2g_{xy}(x, y)dxdy + g_{yy}(x, y)dy^2$$

$$g_{xx}g_{yy} - g_{xy}^2 > 0$$

geodesics

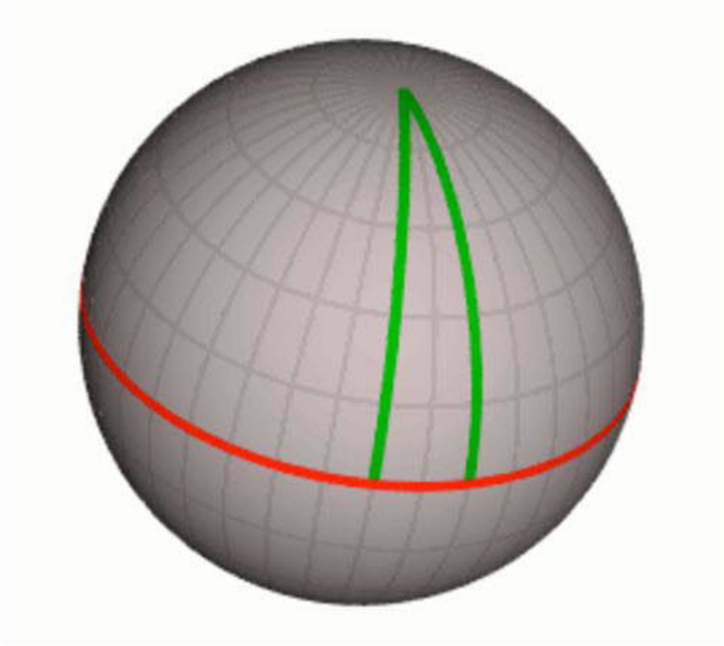


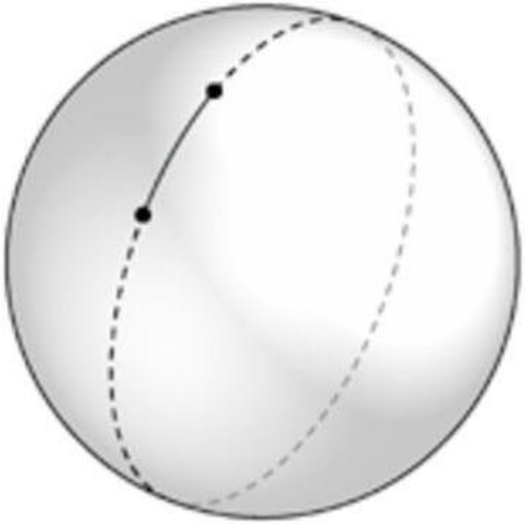


a geodesic is a curve along which the distance between two given points is extremised.

note: important!

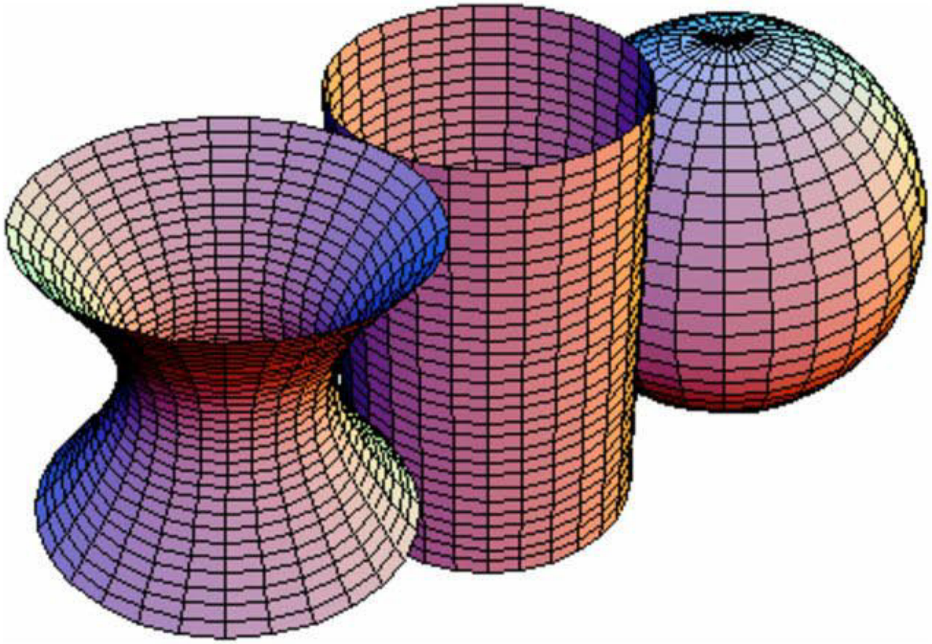
sphere: geodesics





longitudes: yes
latitudes: no

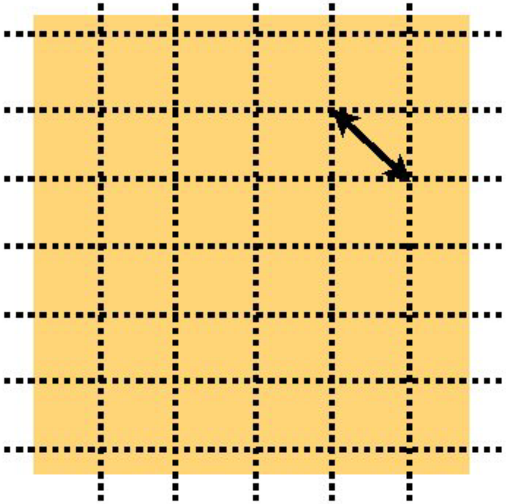
curved ?



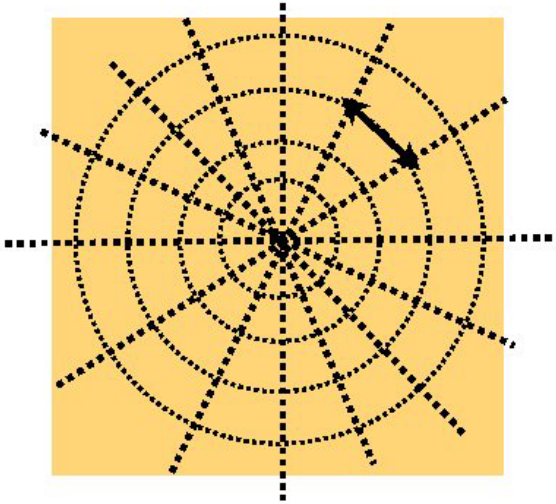
Above image:http://en.wikipedia.org/wiki/Gaussian_curvature

$$ds^2 = dx^2 + dy^2 \qquad ds^2 = g_{xx}(x,y)dx^2 + 2g_{xy}(x,y)dxdy + g_{yy}(x,y)dy^2$$

metric and co-ordinates

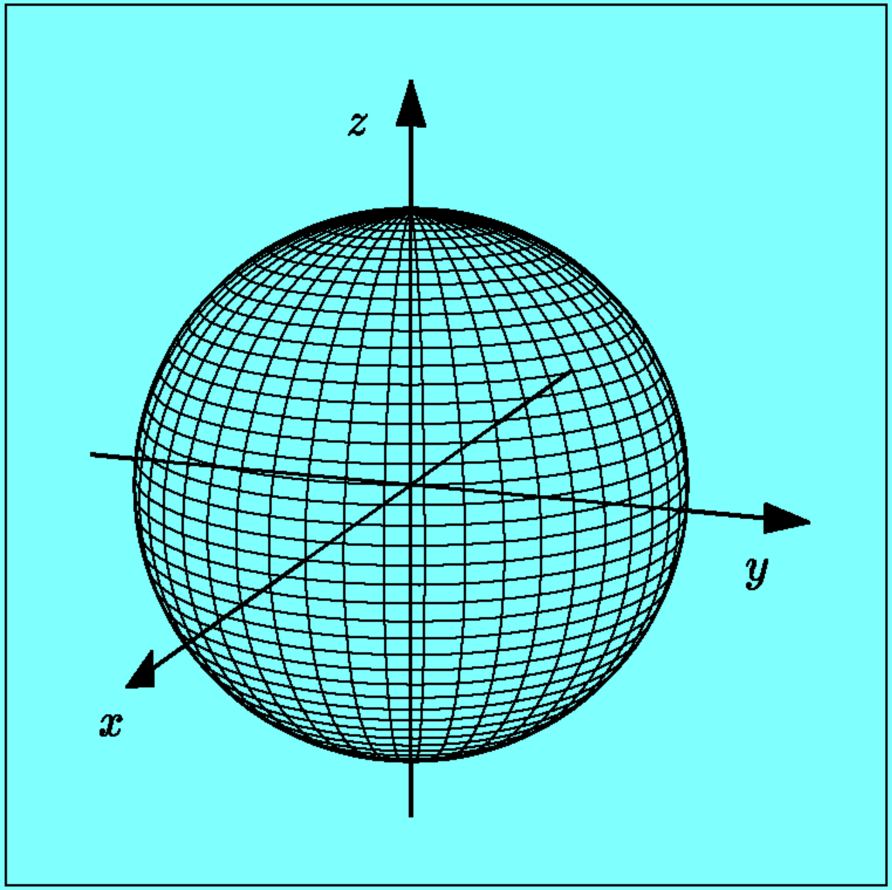


$$ds^2 = dx^2 + dy^2$$



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

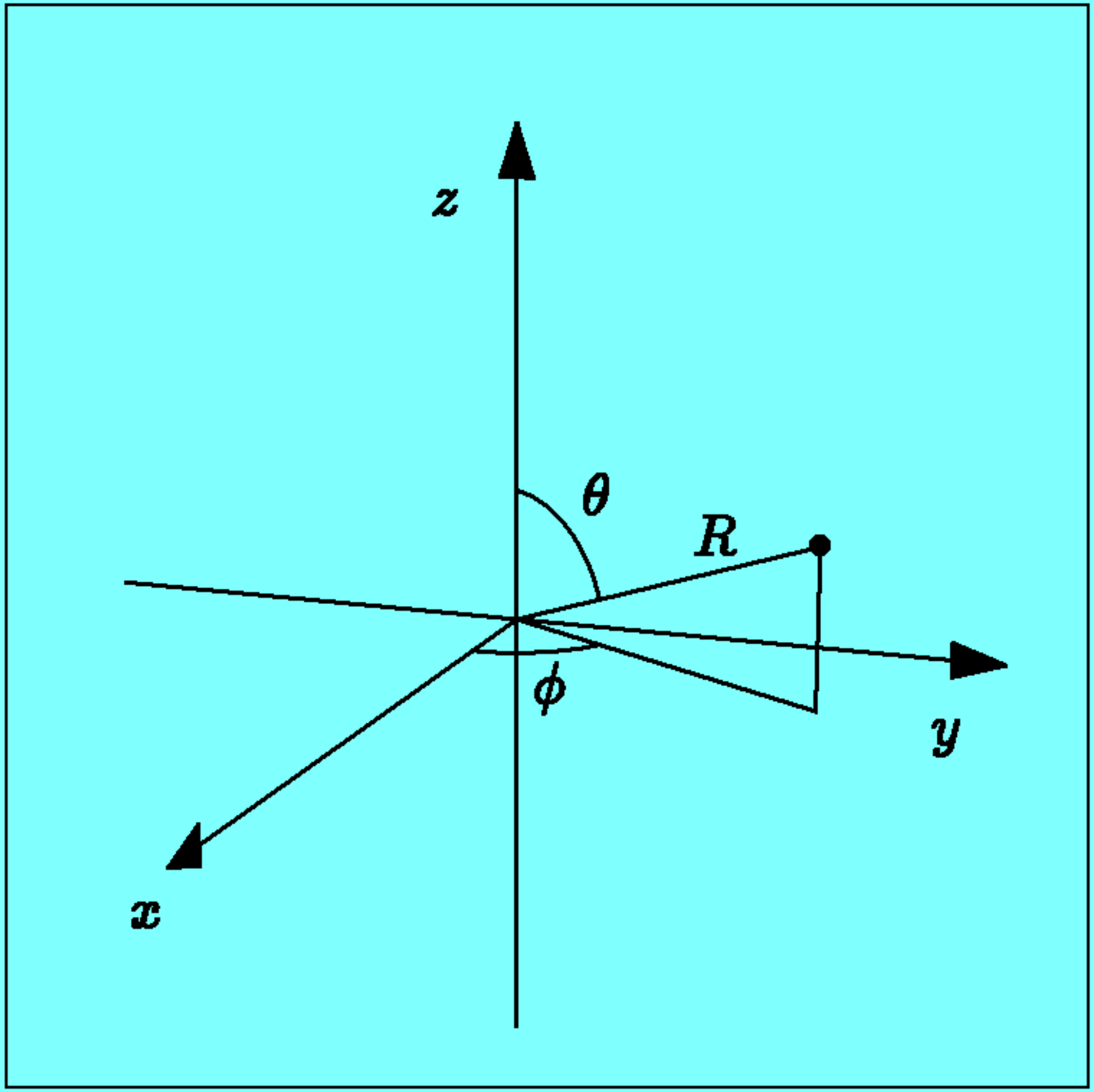
Non-Euclidean Geometry: The Surface of a Sphere



$$x^2 + y^2 + z^2 = R^2 \ .$$

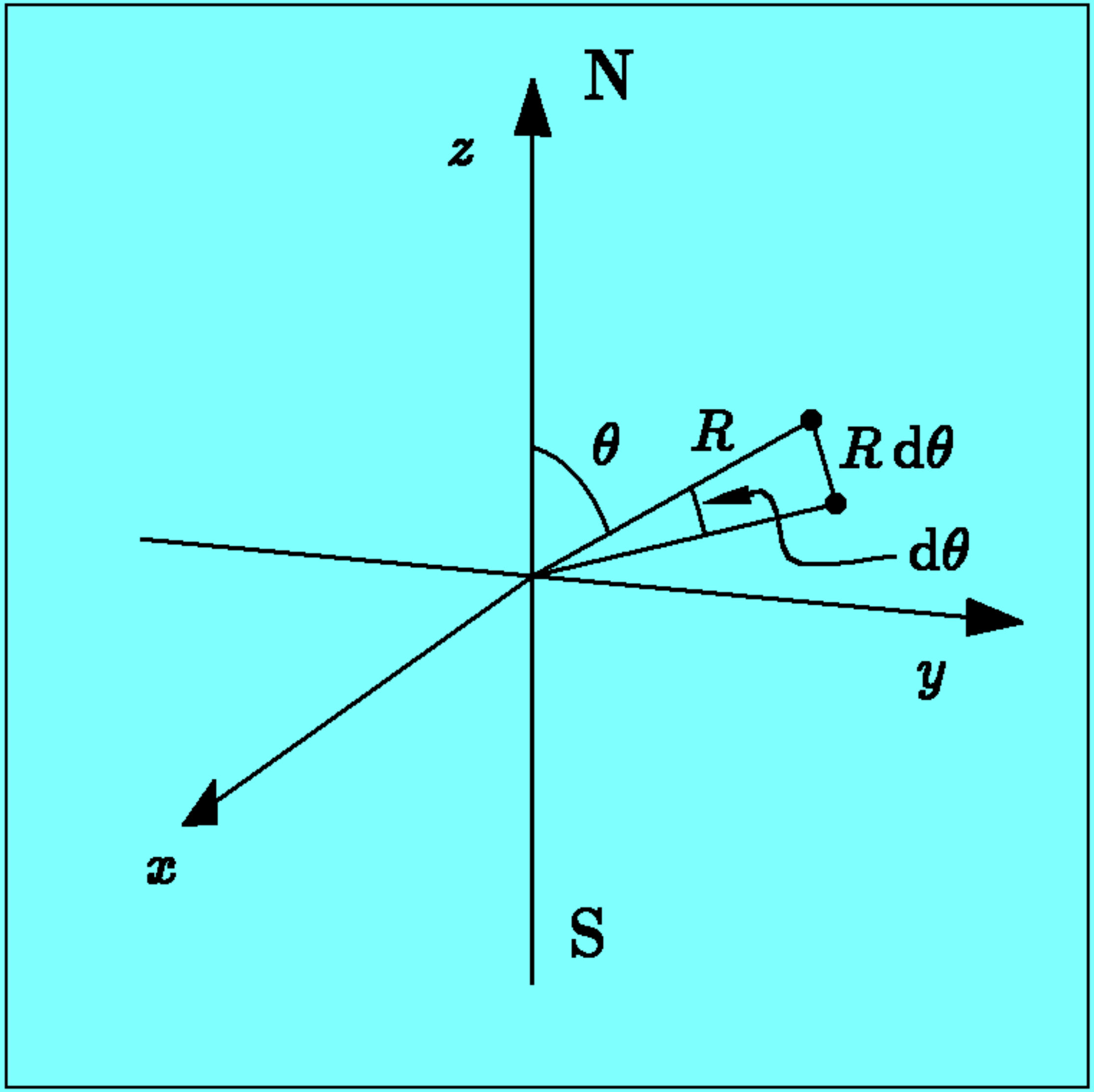
Polar Coordinates:

$$x = R\sin\theta\cos\phi$$
$$y = R\sin\theta\sin\phi$$
$$z = R\cos\theta\ ,$$



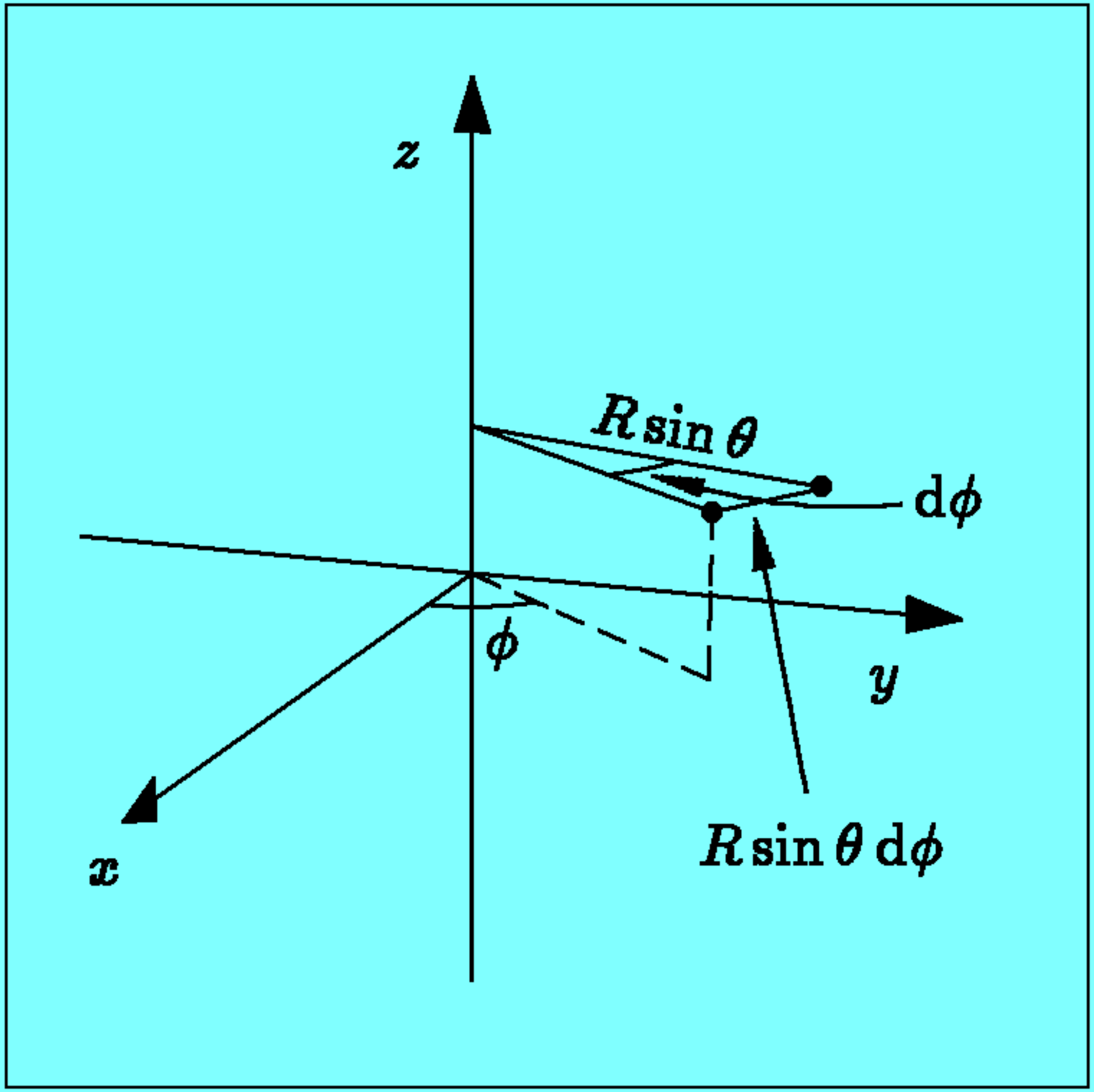
Varying θ:

$$ds = R\,d\theta$$



Varying ϕ:

$$ds = R\sin\theta\, d\phi$$



Varying θ and ϕ

Varying θ: $\quad ds = R\,d\theta$

Varying ϕ: $\quad ds = R\sin\theta\,d\phi$

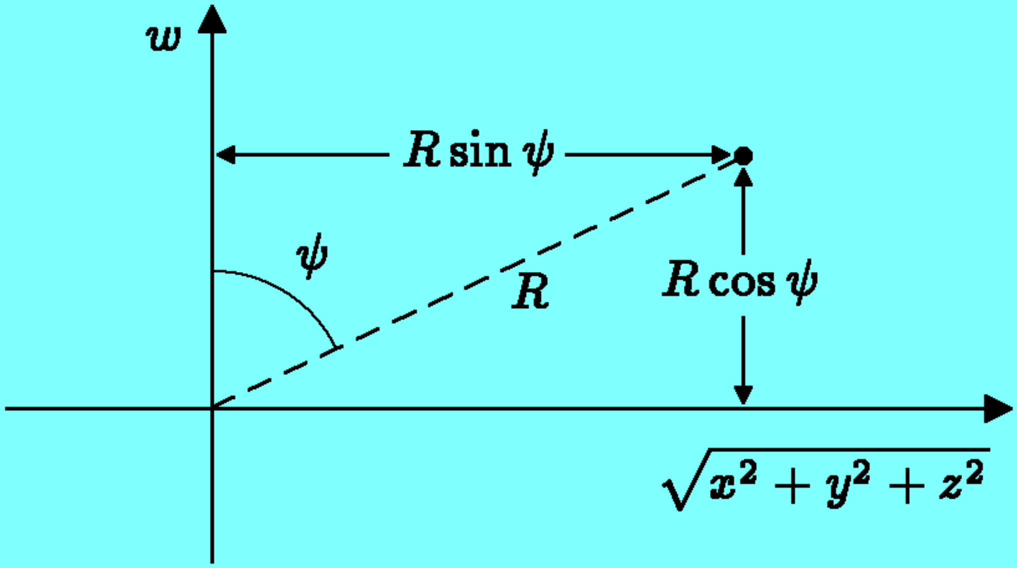
$$ds^2 = R^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)$$

A Closed Three-Dimensional Space

$$\boxed{x^2 + y^2 + z^2 + w^2 = R^2}$$

$$\boxed{\begin{aligned} x &= R \sin\psi \sin\theta \cos\phi \\ y &= R \sin\psi \sin\theta \sin\phi \\ z &= R \sin\psi \cos\theta \\ w &= R \cos\psi \ , \end{aligned}}$$

$$\boxed{ds = R\,d\psi}$$



Metric for the Closed 3D Space

Varying ψ: $\quad ds = R\,d\psi$

Varying θ or ϕ: $\quad ds^2 = R^2 \sin^2\psi (d\theta^2 + \sin^2\theta\, d\phi^2)$

If the variations are orthogonal to each other, then

$$ds^2 = R^2\left[d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right]$$

Proof of Orthogonality of Variations

Let $d\vec{r}_\psi$ = displacement of point when ψ is changed to $\psi + d\psi$.

Let $d\vec{r}_\theta$ = displacement of point when θ is changed to $\theta + d\theta$.

- $d\vec{r}_\theta$ has no w-component $\Longrightarrow$ $d\vec{r}_\psi \cdot d\vec{r}_\theta = d\vec{r}_\psi^{\,(3)} \cdot d\vec{r}_\theta^{\,(3)}$, where (3) denotes the projection into the x-y-z subspace.
- $d\vec{r}_\psi^{\,(3)}$ is radial; $d\vec{r}_\theta^{\,(3)}$ is tangential

$$\Longrightarrow \quad d\vec{r}_\psi^{\,(3)} \cdot d\vec{r}_\theta^{\,(3)} = 0$$

Implications of General Relativity

- $ds^2 = R^2\left[d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right]$, where R is radius of curvature.
- According to GR, matter causes space to curve. So R, the curvature radius, should be determined by the matter.
- From the metric, or from the picture of a sphere of radius R in a 4D Euclidean embedding space, it is clear that R determines the size of the space. But $a(t)$, the scale factor, also determines the size of the space. So they must be proportional.
- But R is in meters, $a(t)$ in meters/notch. So dimensional consistency $\Longrightarrow$ $R \propto a(t)/\sqrt{k}$, since $[k] = \text{notch}^{-2}$.
- In fact,

$$\boxed{R^2(t) = \frac{a^2(t)}{k}\ .}$$

(I do not know any way to explain why the proportionality constant is 1, except by using the full equations of GR.)

- $\mathrm{d}s^2 = R^2\left[\mathrm{d}\psi^2 + \sin^2\psi\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]$, where R is radius of curvature.
- In fact,

$$\boxed{R^2(t) = \frac{a^2(t)}{k}\ .}$$

- So,

$$\boxed{\mathrm{d}s^2 = \frac{a^2(t)}{k}\left[\mathrm{d}\psi^2 + \sin^2\psi\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]\ .}$$

- It is common to introduce a new radial variable $\boxed{r \equiv \sin\psi/\sqrt{k},}$ so $\mathrm{d}r = \cos\psi\,\mathrm{d}\psi/\sqrt{k} = \sqrt{1-kr^2}\,\mathrm{d}\psi/\sqrt{k}$. In terms of r,

$$\boxed{\mathrm{d}s^2 = a^2(t)\left\{\frac{\mathrm{d}r^2}{1-kr^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right\}\ .}$$

This is the spatial part of the Robertson-Walker metric.

Open Universes

☆ For $k > 0$ (closed universe),

$$ds^2 = a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\}$$

describes a homogeneous isotropic universe.

☆ For $k < 0$ (open universe),

$$ds^2 = a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\}$$

still describes a homogeneous isotropic universe.

☆ Properties are very different. The closed universe reaches its equator at $r = 1/\sqrt{k}$, which is a finite distance from the origin,

$$a(t)\int_0^{1/\sqrt{k}} \frac{dr}{\sqrt{1-kr^2}} = \frac{\pi a(t)}{2\sqrt{k}}\ .$$

The total volume is finite. For the open universe, r has no limit, and the volume is infinite.

8.286 Class 11
October 17, 2022

INTRODUCTION TO NON-EUCLIDEAN SPACES, PART 3

Metric for the Closed 3D Space

Varying ψ: $\quad ds = R\,d\psi$

Varying θ or ϕ: $\quad ds^2 = R^2 \sin^2\psi\,(d\theta^2 + \sin^2\theta\, d\phi^2)$

If the variations are orthogonal to each other, then

$$ds^2 = R^2 \left[d\psi^2 + \sin^2\psi \left(d\theta^2 + \sin^2\theta\, d\phi^2 \right) \right]$$

Implications of General Relativity

- $\mathrm{d}s^2 = R^2 \left[\mathrm{d}\psi^2 + \sin^2\psi \left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]$, where R is radius of curvature.
- According to GR, matter causes space to curve. So R, the curvature radius, should be determined by the matter.
- From the metric, or from the picture of a sphere of radius R in a 4D Euclidean embedding space, it is clear that R determines the size of the space. But $a(t)$, the scale factor, also determines the size of the space. So they must be proportional.
- But R is in meters, $a(t)$ in meters/notch. So dimensional consistency $\Longrightarrow$ $R \propto a(t)/\sqrt{k}$, since $[k] = \text{notch}^{-2}$.
- In fact,

$$\boxed{R^2(t) = \frac{a^2(t)}{k}\ .}$$

(I do not know any way to explain why the proportionality constant is 1, except by using the full equations of GR.)

- $ds^2 = R^2 \left[d\psi^2 + \sin^2\psi \left(d\theta^2 + \sin^2\theta \, d\phi^2\right)\right]$, where R is radius of curvature.
- In fact,

$$\boxed{R^2(t) = \frac{a^2(t)}{k} \ .}$$

- So,

$$\boxed{ds^2 = \frac{a^2(t)}{k} \left[d\psi^2 + \sin^2\psi \left(d\theta^2 + \sin^2\theta \, d\phi^2\right)\right] \ .}$$

- It is common to introduce a new radial variable $\boxed{r \equiv \sin\psi/\sqrt{k},}$ so $dr = \cos\psi \, d\psi/\sqrt{k} = \sqrt{1 - kr^2} \, d\psi/\sqrt{k}$. In terms of r,

$$\boxed{ds^2 = a^2(t) \left\{ \frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta \, d\phi^2\right) \right\} \ .}$$

This is the spatial part of the Robertson-Walker metric.

Open Universes

- For $k > 0$ (closed universe),

$$ds^2 = a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\}$$

describes a homogeneous isotropic universe.

- For $k < 0$ (open universe),

$$ds^2 = a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\}$$

still describes a homogeneous isotropic universe.

- Properties are very different. The closed universe reaches its equator at $r = 1/\sqrt{k}$, which is a finite distance from the origin,

$$a(t)\int_0^{1/\sqrt{k}} \frac{dr}{\sqrt{1-kr^2}} = \frac{\pi a(t)}{2\sqrt{k}}\ .$$

The total volume is finite. For the open universe, r has no limit, and the volume is infinite.

From Space to Spacetime

In special relativity,

$$s^2_{AB} \equiv (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 - c^2\,(t_A - t_B)^2 \quad .$$

s^2_{AB} is Lorentz-invariant — it has the same value for all inertial reference frames.

From Space to Spacetime

In special relativity,

$$s^2_{AB} \equiv (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 - c^2\,(t_A - t_B)^2 \ .$$

s^2_{AB} is Lorentz-invariant — it has the same value for all inertial reference frames.

Meaning of s^2_{AB}:

If positive, it is the distance^2 between the two events in the inertial frame in which they are simultaneous. (Spacelike.)

If negative, then $s^2_{AB} = -c^2\,\Delta\tau^2$, where $\Delta\tau$ is the time interval between the two events in the inertial frame in which they occur at the same place. (Timelike.)

If zero, it implies that a light pulse could travel from the earlier to the later event. (Lightlike.)

From Space to Spacetime

In special relativity,

$$s_{AB}^2 \equiv (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 - c^2\,(t_A - t_B)^2 \;.$$

s_{AB}^2 is Lorentz-invariant — it has the same value for all inertial reference frames.

Meaning of s_{AB}^2:

If positive, it is the distance^2 between the two events in the inertial frame in which they are simultaneous. (Spacelike.)

If negative, then $s_{AB}^2 = -c^2\,\Delta\tau^2$, where $\Delta\tau$ is the time interval between the two events in the inertial frame in which they occur at the same place. (Timelike.)

If zero, it implies that a light pulse could travel from the earlier to the later event. (Lightlike.)

☆ If you are interested, Lecture Notes 5 has an appendix which derives the Lorentz transformation from time dilation, Lorentz contraction, and the relativity of simultaneity, and shows that s_{AB}^2 is invariant.

Infinitesimal Separations and the Metric

☆ Following Gauss, we focus on the distance between infinitesimally separated points. So

$$s^2_{AB} \equiv (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2 - c^2\,(t_A - t_B)^2$$

is replaced by

$$\boxed{\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\,\mathrm{d}t^2\ ,}$$

which is called the *Minkowski metric.*

☆ The interpretation is the same as before: $\mathrm{d}s^2 > 0 \implies$ distance^2 in frame where events are simultaneous; $\mathrm{d}s^2 < 0 \implies \mathrm{d}s^2 = -c^2\,\mathrm{d}\tau^2$, where $\mathrm{d}\tau =$ time difference in frame where events are at same place; $\mathrm{d}s^2 = 0 \implies$ light can travel from one event to the other.

☆ This will be our springboard to metric used in general relativity.

Coordinates in Curves Spaces

- ☆ In Newtonian physics or special relativity, coordinates have a direct physical meaning: they directly measure distances or time intervals.
- ☆ In curves spaces, there is generally no way to construct coordinates that are directly connected to distances.
- ☆ For example, on the surface of the Earth we measure East-West position by longitude, but the distance for a longitude distance of 1 degree depends on the latitude.
- ☆ Bottom line: in general relativity (or in any curved space), coordinates are just arbitrary markers, with any set of coordinates in principle as good as any other.
- ☆ Distances are determined from the coordinates, using the metric.
- ☆ If one changes from one coordinate system to another, one changes the metric so that distances remain unchanged.

General Relativity: the Equivalence Principle and Free-Falling Observers

☆ Consider a person holding a rock inside an elevator, initially at rest.

General Relativity: the Equivalence Principle and Free-Falling Observers

☆ Consider a person holding a rock inside an elevator, initially at rest. The person feels the force of gravity pulling down on the rock, and the force of gravity pressing his feet against the floor.

General Relativity: the Equivalence Principle and Free-Falling Observers

- ☆ Consider a person holding a rock inside an elevator, initially at rest. The person feels the force of gravity pulling down on the rock, and the force of gravity pressing his feet against the floor.
- ☆ Now imagine that the elevator cable is cut, so the elevator falls — we assume that there is no friction or air resistance.

General Relativity: the Equivalence Principle and Free-Falling Observers

- ☆ Consider a person holding a rock inside an elevator, initially at rest. The person feels the force of gravity pulling down on the rock, and the force of gravity pressing his feet against the floor.

- ☆ Now imagine that the elevator cable is cut, so the elevator falls — we assume that there is no friction or air resistance. The elevator, person, and rock all accelerate together. The person no longer feels his feet pressed to the floor; if he lets go of the rock, it floats. The effects of gravity have disappeared.

General Relativity: the Equivalence Principle and Free-Falling Observers

- Consider a person holding a rock inside an elevator, initially at rest. The person feels the force of gravity pulling down on the rock, and the force of gravity pressing his feet against the floor.

- Now imagine that the elevator cable is cut, so the elevator falls — we assume that there is no friction or air resistance. The elevator, person, and rock all accelerate together. The person no longer feels his feet pressed to the floor; if he lets go of the rock, it floats. The effects of gravity have disappeared.

- The *Equivalence Principle* says that the disappearance of gravity is precise: as long as the elevator is small enough so that the gravitational field is uniform, then there is absolutely no way that the person in the free-falling elevator can detect the gravitational field of the Earth.

- ☆ The *Equivalence Principle* says that the disappearance of gravity is precise: as long as the elevator is small enough so that the gravitational field is uniform, then there is absolutely no way that the person in the free-falling elevator can detect the gravitational field of the Earth.

- ☆ The person in the elevator is called a *free-falling observer*, and the local coordinate system that he would construct in his immediate vicinity is called a free-falling coordinate system. The metric for the free-falling coordinates, in the immediate vicinity of the person, is described by the Minkowski metric. It is called *locally Minkowskian*.

- ☆ We mentioned earlier that any quadratic metric for space (i.e., a positive definite metric) is locally Euclidean. If the metric is negative for one direction, then it is always locally Minkowskian.

- ☆ Not as simple as it sounds! If you calculate the bending of a light beam by gravity taking into account only the acceleration of the elevator, you will get only half the GR answer. The correct free-falling coordinate system is not just an accelerating version of a Euclidean coordinate system, but also takes into account the bending of space caused by gravity (in GR).

Adding Time to the Robertson–Walker Metric

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} .$$

Why does dt^2 term look like it does:

- The coefficient of dt^2 term must be independent of position, due to homogeneity.
- Terms such as $dt\,dr$ or $dt\,d\phi$ cannot appear, due to isotropy. That is, a term $dt\,dr$ would behave differently for $dr > 0$ and $dr < 0$, creating an asymmetry between the $+r$ and -r directions.
- The coefficient must be negative, to match the sign in Minkowski space for a locally free-falling coordinate system.

Adding Time to the Robertson-Walker Metric

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}t^2 + a^2(t)\left\{\frac{\mathrm{d}r^2}{1-kr^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right\} .$$

Meaning:

- If $ds^2 > 0$, it is the square of the spatial separation measured by a local free-falling observer for whom the two events happen at the same time.
- If $ds^2 < 0$, it is $-c^2$ times the square of the time separation measured by a local free-falling observer for whom the two events happen at the same location.
- If $ds^2 = 0$, then the two events can be joined by a light pulse.

Summary: Metrics of Interest

Minkowski Metric: (Special relativity)

$$\begin{aligned} ds^2 &= -c^2 dt^2 + dx^2 + dy^2 + dz^2 \\ &= -c^2 dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2) \ . \end{aligned}$$

Robertson–Walker Metric:

$$ds^2 = -c^2\, dt^2 + a^2(t) \left\{ \frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta \, d\phi^2 \right) \right\} \ .$$

Meaning: If $ds^2 > 0$, ds is distance in freely falling frame in which events are simultaneous. If $ds^2 < 0$, $ds^2 = -c^2 d\tau^2$, where $d\tau$ is time interval in freely falling frame in which events occur at same point. If $ds^2 = 0$, events are lightlike separated.

Geodesics in General Relativity

A geodesic is a path connecting two points in spacetime, with the property that the length of the curve is stationary with respect to small changes in the path. It can be a maximum, minimum, or saddle point.

In a curved spacetime, a geodesic is the closest thing to a straight line that exists.

In general relativity, if no forces act on a particle other than gravity, the particle travels on a geodesic.

8.286 Class 12
October 19, 2022

INTRODUCTION TO NON-EUCLIDEAN SPACES, PART 4

Summary: Metrics of Interest

Minkowski Metric: (Special relativity)

$$\begin{aligned} ds^2 &= -c^2 dt^2 + dx^2 + dy^2 + dz^2 \\ &= -c^2 dt^2 + dr^2 + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2) \ . \end{aligned}$$

Robertson–Walker Metric:

$$ds^2 = -c^2 \, dt^2 + a^2(t) \left\{ \frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta \, d\phi^2 \right) \right\} \ .$$

Meaning: If $ds^2 > 0$, ds is distance in freely falling frame in which events are simultaneous. If $ds^2 < 0$, $ds^2 = -c^2 d\tau^2$, where $d\tau$ is time interval in freely falling frame in which events occur at same point. If $ds^2 = 0$, events are lightlike separated.

Geodesics in General Relativity

A geodesic is a path connecting two points in spacetime, with the property that the length of the curve is stationary with respect to small changes in the path. It can be a maximum, minimum, or saddle point.

In a curved spacetime, a geodesic is the closest thing to a straight line that exists.

In general relativity, if no forces act on a particle other than gravity, the particle travels on a geodesic.

Geodesics in Two Spatial Dimensions

Metric:

$$ds^2 = g_{xx}dx^2 + g_{xy}dx\,dy + g_{yx}dy\,dx + g_{yy}dy^2 \ .$$

Let $x^1 \equiv x$, $x^2 \equiv y$, so x^i is either, as $i = 1$ or 2.

$$ds^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij}(x^\ell)\,dx^i\,dx^j$$

$$= g_{ij}(x^\ell)\,dx^i\,dx^j \ .$$

Einstein summation convention: repeated indices within one term are summed over coordinate indices (1 and 2), unless otherwise specified.

The sum is always over one upper index and one lower, but we will not discuss why some indices are written as upper and some as lower.

$g_{ij}\left(x^\ell\right)$ indicates that g_{ij} is a function of all the components of x^ℓ. I.e., when x^ℓ occurs as an argument of a function, it is shorthand for (x^1, x^2). By contrast, dx^i denotes the i'th component of dx, meaning dx^1 if $i = 1$, or dx^2 if $i = 2$.

The Length of Path

Consider a path from A to B.

Path description: $x^i(\lambda)$, where λ is parameter running from 0 to λ_f.

$$x^i(0) = x^i_A, \qquad x^i(\lambda_f) = x^i_B \ .$$

Between λ and $\lambda + \mathrm{d}\lambda$,

$$\mathrm{d}x^i = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\mathrm{d}\lambda \ ,$$

so

$$\mathrm{d}s^2 = \boldsymbol{g_{ij}(x^\ell)\,\mathrm{d}x^i\,\mathrm{d}x^j} = g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\,\mathrm{d}\lambda^2 \ ,$$

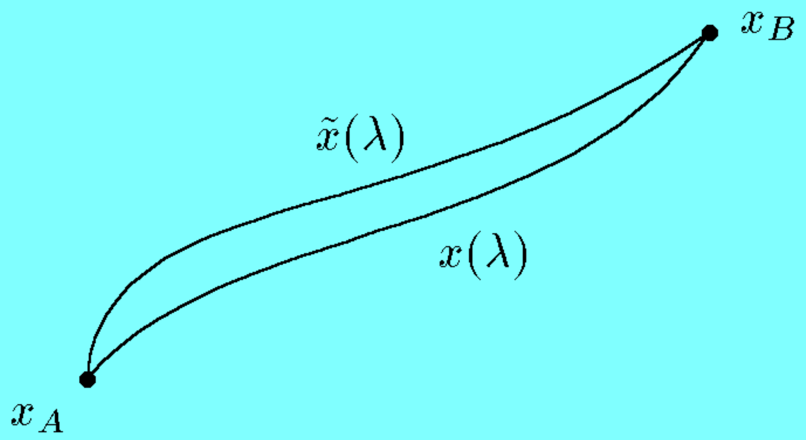
and then

$$\mathrm{d}s = \sqrt{g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda \ ,$$

and

$$\boxed{S[x^i(\lambda)] = \int_0^{\lambda_f}\sqrt{g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda \ .}$$

Varying the Path



$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \ ,$$

where

$$w^i(0) = 0 \ , \qquad w^i(\lambda_f) = 0 \ .$$

Geodesic condition:

$$\boxed{\left.\frac{\mathrm{d}\, S\left[\tilde{x}^i(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = 0 \qquad \text{for all } w^i(\lambda) \quad .}$$

$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \; .$$

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{g_{ij}\big(\tilde{x}^\ell(\lambda)\big)\,\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda \; .$$

Define

$$A(\lambda,\alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda} \; ,$$

so we can write

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{A(\lambda,\alpha)}\,\mathrm{d}\lambda \; .$$

Using chain rule,

$$\frac{\mathrm{d}f\big(x(\alpha),y(\alpha)\big)}{\mathrm{d}\alpha} = \frac{\partial f(x,y)}{\partial x}\frac{\mathrm{d}x(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f(x,y)}{\partial y}\frac{\mathrm{d}y(\alpha)}{\mathrm{d}\alpha} = \frac{\partial f(x^\ell)}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}\alpha} \; ,$$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}g_{ij}\big(\tilde{x}^\ell(\lambda)\big)\bigg|_{\alpha=0} = \left[\frac{\partial g_{ij}}{\partial\tilde{x}^k}\frac{\partial\tilde{x}^k}{\partial\alpha}\right]_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\big(x^\ell(\lambda)\big)\;\frac{\partial\tilde{x}^k}{\partial\alpha}\bigg|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\big(x^\ell(\lambda)\big)w^k \; ,$$

$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \ .$$

$$A(\lambda, \alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right) \frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda} \frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda} \ .$$

Using chain rule, $\dfrac{\mathrm{d}f\big(x(\alpha), y(\alpha)\big)}{\mathrm{d}\alpha} = \dfrac{\partial f(x,y)}{\partial x}\dfrac{\mathrm{d}x(\alpha)}{\mathrm{d}\alpha} + \dfrac{\partial f(x,y)}{\partial y}\dfrac{\mathrm{d}y(\alpha)}{\mathrm{d}\alpha}$,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\bigg|_{\alpha=0} = \left[\frac{\partial g_{ij}}{\partial \tilde{x}^k}\frac{\partial \tilde{x}^k}{\partial \alpha}\right]_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\big(x^\ell(\lambda)\big) \left.\frac{\partial \tilde{x}^k}{\partial \alpha}\right|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\big(x^\ell(\lambda)\big) w^k \ .$$

Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\right) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\left[\frac{\mathrm{d}x^i(\lambda)}{\mathrm{d}\lambda} + \alpha \frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda}\right] = \frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda} \ .$$

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{A(\lambda,\alpha)}\,\mathrm{d}\lambda \ ,$$

where

$$A(\lambda,\alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda} \ ,$$

with

$$\left.\frac{\mathrm{d}}{\mathrm{d}\alpha} g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\right|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right) w^k \ , \qquad \frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\right) = \frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda} \ .$$

Then

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^\ell(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k} w^k \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + \right.$$

$$\left. + g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}w^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda \ ,$$

where the metric g_{ij} is to be evaluated at $x^\ell(\lambda)$.

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0}=\frac{1}{2}\int_0^{\lambda_f}\frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k}w^k\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}+\right.$$
$$\left.+g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}+g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}w^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

Manipulating "dummy" indices: in third term, replace $i \to j$ and $j \to i$, and recall that $g_{ij} = g_{ji}$. Then 2nd & 3rd term are equal:

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0}=\frac{1}{2}\int_0^{\lambda_f}\frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k}w^k\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}+2g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

Repeating,

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^\ell(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k} w^k \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + 2g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda \ .$$

Integration by Parts: Integral depends on both w^k and $\mathrm{d}w^i/\mathrm{d}\lambda$. Can eliminate $\mathrm{d}w^i/\mathrm{d}\lambda$ by integrating by parts:

$$\int_0^{\lambda_f}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right]\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\,\mathrm{d}\lambda = \int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^i\right]\mathrm{d}\lambda$$
$$- \int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] w^i\,\mathrm{d}\lambda \ .$$

But

$$\int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^i\right]\mathrm{d}\lambda = \left.\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^i\right]\right|_{\lambda=0}^{\lambda=\lambda_f} = 0 \ ,$$

since $w^i(\lambda)$ vanishes at $\lambda = 0$ and $\lambda = \lambda_f$.

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k} w^k \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + 2 g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \left\{\frac{1}{\sqrt{A}}\frac{\partial g_{ij}}{\partial x^k}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^k - 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] w^i\right\}\mathrm{d}\lambda\ .$$

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \left\{ \frac{1}{\sqrt{A}}\frac{\partial g_{ij}}{\partial x^k}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^k - 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] w^i \right\} \mathrm{d}\lambda \ .$$

Complication: one term is proportional to w^k, and the other is proportional to w^i. But with more index juggling, we can fix that. In 1st term replace $i \to j, j \to k, k \to i$:

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \int_0^{\lambda_f} \left\{ \frac{1}{2\sqrt{A}}\frac{\partial g_{jk}}{\partial x^i}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda} - \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] \right\} w^i(\lambda)\,\mathrm{d}\lambda \ .$$

To vanish ***for all*** $\boldsymbol{w^i(\lambda)}$ which vanish at $\lambda = 0$ and $\lambda = \lambda_f$, the quantity in curly brackets must vanish. If not, then suppose that $\{\,\}_i > 0$ for some $i = i_0$ and for some $\lambda = \lambda_0$. By continuity, $\{\,\}_{i_0} > 0$ in some neighborhood of λ_0. Choose $w^{i_0}(\lambda)$ to be positive in this neighborhood, and zero everywhere else, with $w^j(\lambda) \equiv 0$ for $j \neq i_0$, and one has a contradiction.

So

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] = \frac{1}{2\sqrt{A}}\frac{\partial g_{jk}}{\partial x^i}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \quad .}$$

Repeating,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] = \frac{1}{2\sqrt{A}} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \quad .}$$

This is ***complicated,*** since A is complicated.

Simplify by choice of parameterization: This result is valid for any parameterization. We don't need that! We can choose λ to be the path length. Since

$$\mathrm{d}s = \sqrt{g_{ij}\left(x^\ell(\lambda)\right) \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda = \sqrt{A}\,\mathrm{d}\lambda \ ,$$

we see that $\mathrm{d}\lambda = \mathrm{d}s$ implies

$$A = 1 \qquad (\text{for } \lambda = \text{path length}).$$

Then

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}s}\left[g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}s}\right] = \frac{1}{2} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}s} \frac{\mathrm{d}x^k}{\mathrm{d}s} \ .}$$

8.286 Class 13
October 24, 2022

INTRODUCTION TO NON-EUCLIDEAN SPACES, PART 5

Geodesics in General Relativity

A geodesic is a path connecting two points in spacetime, with the property that the length of the curve is stationary with respect to small changes in the path. It can be a maximum, minimum, or saddle point.

In a curved spacetime, a geodesic is the closest thing to a straight line that exists.

In general relativity, if no forces act on a particle other than gravity, the particle travels on a geodesic.

Geodesics in Two Spatial Dimensions

Metric:

$$\mathrm{d}s^2 = g_{xx}\mathrm{d}x^2 + g_{xy}\mathrm{d}x\,\mathrm{d}y + g_{yx}\mathrm{d}y\,\mathrm{d}x + g_{yy}\mathrm{d}y^2 \ .$$

Let $x^1 \equiv x$, $x^2 \equiv y$, so x^i is either, as $i = 1$ or 2.

$$\mathrm{d}s^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij}(x^\ell)\,\mathrm{d}x^i\,\mathrm{d}x^j$$

$$= g_{ij}(x^\ell)\,\mathrm{d}x^i\,\mathrm{d}x^j \ .$$

Einstein summation convention: repeated indices within one term are summed over coordinate indices (1 and 2), unless otherwise specified.

The sum is always over one upper index and one lower, but we will not discuss why some indices are written as upper and some as lower.

$g_{ij}\left(x^\ell\right)$ indicates that g_{ij} is a function of all the components of x^ℓ. I.e., when x^ℓ occurs as an argument of a function, it is shorthand for (x^1, x^2). By contrast, dx^i denotes the i'th component of dx, meaning dx^1 if $i = 1$, or dx^2 if $i = 2$.

The Length of Path

Consider a path from A to B.

Path description: $x^i(\lambda)$, where λ is parameter running from 0 to λ_f.

$$x^i(0) = x^i_A, \qquad x^i(\lambda_f) = x^i_B \ .$$

Between λ and $\lambda + \mathrm{d}\lambda$,

$$\mathrm{d}x^i = \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\mathrm{d}\lambda \ ,$$

so

$$\mathrm{d}s^2 = \boldsymbol{g_{ij}(x^\ell)\,\mathrm{d}x^i\,\mathrm{d}x^j} = g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\,\mathrm{d}\lambda^2 \ ,$$

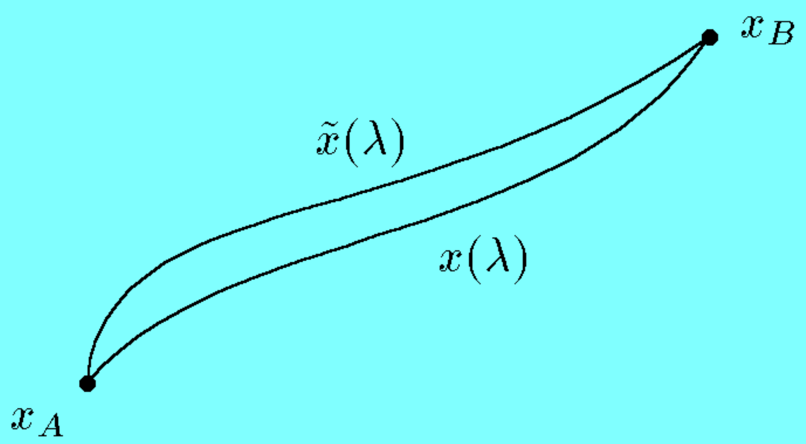
and then

$$\mathrm{d}s = \sqrt{g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda \ ,$$

and

$$\boxed{S[x^i(\lambda)] = \int_0^{\lambda_f}\sqrt{g_{ij}\left(x^\ell(\lambda)\right)\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\,\mathrm{d}\lambda \ .}$$

Varying the Path



$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \ ,$$

where

$$w^i(0) = 0 \ , \qquad w^i(\lambda_f) = 0 \ .$$

Geodesic condition:

$$\boxed{\left. \frac{\mathrm{d}\, S\left[\tilde{x}^i(\lambda)\right]}{\mathrm{d}\alpha} \right|_{\alpha=0} = 0 \qquad \text{for all } w^i(\lambda) \quad .}$$

$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \ .$$

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{g_{ij}\left(\tilde{x}^\ell(\lambda)\right) \frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda} \frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda}}\, \mathrm{d}\lambda \ .$$

Define

$$A(\lambda, \alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right) \frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda} \frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda} \ ,$$

so we can write

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{A(\lambda, \alpha)}\, \mathrm{d}\lambda \ .$$

Using chain rule,

$$\frac{\mathrm{d}f\left(x(\alpha), y(\alpha)\right)}{\mathrm{d}\alpha} = \frac{\partial f(x,y)}{\partial x} \frac{\mathrm{d}x(\alpha)}{\mathrm{d}\alpha} + \frac{\partial f(x,y)}{\partial y} \frac{\mathrm{d}y(\alpha)}{\mathrm{d}\alpha} = \frac{\partial f(x^\ell)}{\partial x^i} \frac{\mathrm{d}x^i}{\mathrm{d}\alpha} \ ,$$

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\Big|_{\alpha=0} = \left[\frac{\partial g_{ij}}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^k}{\partial \alpha}\right]_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right) \left.\frac{\partial \tilde{x}^k}{\partial \alpha}\right|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right) w^k \ ,$$

$$\tilde{x}^i(\lambda) = x^i(\lambda) + \alpha w^i(\lambda) \ .$$

$$A(\lambda,\alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right) \frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda} \ .$$

Using chain rule, $\dfrac{\mathrm{d}f\left(x(\alpha),y(\alpha)\right)}{\mathrm{d}\alpha} = \dfrac{\partial f(x,y)}{\partial x}\dfrac{\mathrm{d}x(\alpha)}{\mathrm{d}\alpha} + \dfrac{\partial f(x,y)}{\partial y}\dfrac{\mathrm{d}y(\alpha)}{\mathrm{d}\alpha}$,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\bigg|_{\alpha=0} = \left[\frac{\partial g_{ij}}{\partial \tilde{x}^k}\frac{\partial \tilde{x}^k}{\partial \alpha}\right]_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right) \left.\frac{\partial \tilde{x}^k}{\partial \alpha}\right|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right) w^k \ .$$

Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\right) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\left[\frac{\mathrm{d}x^i(\lambda)}{\mathrm{d}\lambda} + \alpha\frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda}\right] = \frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda} \ .$$

$$S\left[\tilde{x}^i(\lambda)\right] = \int_0^{\lambda_f} \sqrt{A(\lambda,\alpha)}\,\mathrm{d}\lambda\ ,$$

where

$$A(\lambda,\alpha) = g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\frac{\mathrm{d}\tilde{x}^j}{\mathrm{d}\lambda}\ ,$$

with

$$\left.\frac{\mathrm{d}}{\mathrm{d}\alpha}g_{ij}\left(\tilde{x}^\ell(\lambda)\right)\right|_{\alpha=0} = \frac{\partial g_{ij}}{\partial x^k}\left(x^\ell(\lambda)\right)w^k\ , \qquad \frac{\mathrm{d}}{\mathrm{d}\alpha}\left(\frac{\mathrm{d}\tilde{x}^i}{\mathrm{d}\lambda}\right) = \frac{\mathrm{d}w^i(\lambda)}{\mathrm{d}\lambda}\ .$$

Then

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^\ell(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f}\frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k}w^k\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + \right.$$

$$\left. + g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}w^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ ,$$

where the metric g_{ij} is to be evaluated at $x^\ell(\lambda)$.

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k}w^k\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + \right.$$

$$\left. + g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}w^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

Manipulating "dummy" indices: in third term, replace $i \to j$ and $j \to i$, and recall that $g_{ij} = g_{ji}$. Then 2nd & 3rd term are equal:

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k}w^k\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + 2g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

Repeating,

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k} w^k \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + 2g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

Integration by Parts: Integral depends on both w^k and $\mathrm{d}w^i/\mathrm{d}\lambda$. Can eliminate $\mathrm{d}w^i/\mathrm{d}\lambda$ by integrating by parts:

$$\int_0^{\lambda_f}\left[\frac{1}{\sqrt{A}}g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right]\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\,\mathrm{d}\lambda = \int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}}g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}w^i\right]\mathrm{d}\lambda$$
$$-\int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}}g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right]w^i\,\mathrm{d}\lambda\ .$$

But

$$\int_0^{\lambda_f}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}}g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}w^i\right]\mathrm{d}\lambda = \left.\left[\frac{1}{\sqrt{A}}g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}w^i\right]\right|_{\lambda=0}^{\lambda=\lambda_f} = 0\ ,$$

since $w^i(\lambda)$ vanishes at $\lambda = 0$ and $\lambda = \lambda_f$.

Review from the previous lecture

$$\left.\frac{\mathrm{d}S\left[\tilde{x}^{\ell}(\lambda)\right]}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \frac{1}{\sqrt{A(\lambda,0)}}\left\{\frac{\partial g_{ij}}{\partial x^k} w^k \frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} + 2g_{ij}\frac{\mathrm{d}w^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right\}\mathrm{d}\lambda\ .$$

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f}\left\{\frac{1}{\sqrt{A}}\frac{\partial g_{ij}}{\partial x^k}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^k - 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] w^i\right\}\mathrm{d}\lambda\ .$$

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \frac{1}{2}\int_0^{\lambda_f} \left\{ \frac{1}{\sqrt{A}} \frac{\partial g_{ij}}{\partial x^k} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} w^k - 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] w^i \right\} \mathrm{d}\lambda \; .$$

Complication: one term is proportional to w^k, and the other is proportional to w^i. But with more index juggling, we can fix that. In 1st term replace $i \to j, j \to k, k \to i$:

$$\left.\frac{\mathrm{d}S}{\mathrm{d}\alpha}\right|_{\alpha=0} = \int_0^{\lambda_f} \left\{ \frac{1}{2\sqrt{A}} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} - \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] \right\} w^i(\lambda)\, \mathrm{d}\lambda \; .$$

To vanish ***for all*** $\boldsymbol{w^i(\lambda)}$ which vanish at $\lambda = 0$ and $\lambda = \lambda_f$, the quantity in curly brackets must vanish. If not, then suppose that $\{\,\}_i > 0$ for some $i = i_0$ and for some $\lambda = \lambda_0$. By continuity, $\{\,\}_{i_0} > 0$ in some neighborhood of λ_0. Choose $w^{i_0}(\lambda)$ to be positive in this neighborhood, and zero everywhere else, with $w^j(\lambda) \equiv 0$ for $j \neq i_0$, and one has a contradiction.

So

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] = \frac{1}{2\sqrt{A}} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \quad .}$$

Repeating,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\frac{1}{\sqrt{A}} g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right] = \frac{1}{2\sqrt{A}} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \quad .}$$

This is ***complicated,*** since A is complicated.

Simplify by choice of parameterization: This result is valid for any parameterization. We don't need that! We can choose λ to be the path length. Since

$$\mathrm{d}s = \sqrt{g_{ij}\left(x^\ell(\lambda)\right) \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\, \mathrm{d}\lambda = \sqrt{A}\, \mathrm{d}\lambda \ ,$$

we see that $\mathrm{d}\lambda = \mathrm{d}s$ implies

$$A = 1 \qquad (\text{for } \lambda = \text{path length}).$$

Then

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}s}\left[g_{ij} \frac{\mathrm{d}x^j}{\mathrm{d}s}\right] = \frac{1}{2} \frac{\partial g_{jk}}{\partial x^i} \frac{\mathrm{d}x^j}{\mathrm{d}s} \frac{\mathrm{d}x^k}{\mathrm{d}s} \ .}$$

Alternative Form of Geodesic Equation

Most books write the geodesic equation differently, as

$$\boxed{\frac{\mathrm{d}^2 x^i}{\mathrm{d}s^2} = -\Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s}\ ,}$$

where

$$\boxed{\Gamma^i_{jk} = \frac{1}{2}g^{i\ell}\left(\partial_j g_{\ell k} + \partial_k g_{\ell j} - \partial_\ell g_{jk}\right)}$$

and $g^{i\ell}$ is the matrix inverse of g_{ij}. The quantity Γ^i_{jk} is called the affine connection.

If you are interested, see the lecture notes. If you are not interested, you can skip this.

BLACK HOLES (Fun!)

The Schwarzschild Metric:

For any spherically symmetric distribution of mass, outside the mass the metric is given by the Schwarzschild metric,

$$\mathrm{d}s^2 = -c^2\mathrm{d}\tau^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2\mathrm{d}t^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\ ,$$

where M is the total mass, G is Newton's gravitational constant, c is the speed of light, and θ and ϕ have the usual polar-angle ranges.

Schwarzschild Horizon

$$ds^2 = -c^2 d\tau^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2) \; .$$

The metric is singular at

$$r = R_S \equiv \frac{2GM}{c^2} \; ,$$

where the coefficient of $c^2 dt^2$ vanishes, and the coefficient of dr^2 is infinite.

Surprisingly, this singularity is not real — it is a coordinate artifact. There are other coordinate systems where the metric is smooth at R_S.

But R_S is a ***horizon:*** If you fall past the horizon, there is no return, even if you are photon.

Schwarzschild Radius of the Sun

$$R_{S,\odot} = \frac{2GM}{c^2}$$

$$= \frac{2 \times 6.673 \times 10^{-11} \text{ m}^3\text{-kg}^{-1}\text{-s}^{-2} \times 1.989 \times 10^{30} \text{ kg}}{(2.998 \times 10^8 \text{ m-s}^{-1})^2}$$

$$= 2.95 \text{ km} .$$

- ☆ If the Sun were compressed to this radius, it would become a black hole. Since the Sun is much larger than R_S, and the Schwarzschild metric is only valid outside the matter, there is no Schwarzschild horizon in the Sun.
- ☆ At the center of our galaxy is a supermassive black hole, with $M = 4.1 \times 10^6 \, M_\odot$. This gives $R_S = 1.2 \times 10^{10}$ meters $\approx 1/4$ of radius of orbit of Mercury ≈ 17 times radius of Sun.

Radial Geodesics in the Schwarzschild Metric

$$\mathrm{d}s^2 = -c^2\mathrm{d}\tau^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2\mathrm{d}t^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} \mathrm{d}r^2$$
$$+ r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2) \ .$$

Consider a particle released from rest at $r = r_0$.

r is a "radial coordinate," but not the radius, since it is not the distance from some center. If r is varied by $\mathrm{d}r$, the distance traveled is not $\mathrm{d}r$, but $\mathrm{d}r/\sqrt{1 - 2GM/rc^2}$. r can be called the "circumferential radius," since the term $r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)$ in the metric implies that the circumference of a circle about the origin is $2\pi r$.

By symmetry, the particle will fall straight down, with no change in θ or ϕ. Spherical symmetry implies that all directions in θ and ϕ are equivalent, so any motion in θ–ϕ space would violate this symmetry.

Particle Trajectories in Spacetime

Particle trajectories are timelike, so we use proper time τ to parameterize them, where $\mathrm{d}s^2 \equiv -c^2\mathrm{d}\tau^2$. This implies that $A = -c^2$, instead of $A = 1$, but as long as A is constant, it drops out of the geodesic equation.

By tradition, the spacetime indices in general relativity are denoted by Greek letters such as μ, ν, λ, σ, and are summed from 0 to 3, where $x^0 \equiv t$.

The geodesic equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\left[g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}s}\right] = \frac{1}{2}\frac{\partial g_{jk}}{\partial x^i}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s}$$

is then rewritten as

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right] = \frac{1}{2}\frac{\partial g_{\lambda\sigma}}{\partial x^\mu}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\ .}$$

Radial Trajectory Equations

Only $\mathrm{d}r/\mathrm{d}\tau$ and $\mathrm{d}t/\mathrm{d}\tau$ are nonzero. But they are related by the metric:

$$c^2 \mathrm{d}\tau^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2\, \mathrm{d}t^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \mathrm{d}r^2$$

implies that

$$c^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2 \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 \; .$$

Then, looking at the $\mu = r$ geodesic equation,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g_{\mu\nu}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right] = \frac{1}{2}\,\frac{\partial g_{\lambda\sigma}}{\partial x^\mu}\,\frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\,\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}$$

implies that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g_{rr}\frac{\mathrm{d}r}{\mathrm{d}\tau}\right] = \frac{1}{2}\partial_r g_{rr}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + \frac{1}{2}\partial_r g_{tt}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \; ,$$

where

$$g_{rr} = \left(1 - \frac{2GM}{rc^2}\right)^{-1} \; , \quad g_{tt} = -c^2\left(1 - \frac{2GM}{rc^2}\right) \; .$$

Repeating,

$$c^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2 \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 \; .$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g_{rr}\frac{\mathrm{d}r}{\mathrm{d}\tau}\right] = \frac{1}{2}\partial_r g_{rr}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + \frac{1}{2}\partial_r g_{tt}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \; ,$$

where

$$g_{rr} = \left(1 - \frac{2GM}{rc^2}\right)^{-1} \; , \quad g_{tt} = -c^2\left(1 - \frac{2GM}{rc^2}\right) \; .$$

Expand

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[g_{rr}\frac{\mathrm{d}r}{\mathrm{d}\tau}\right]$$

with the product rule, replace $(\mathrm{d}t/\mathrm{d}\tau)^2$ using the equation above, and simplify. Result:

$$\boxed{\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} = -\frac{GM}{r^2} \; ,}$$

which looks just like Newton, but it is not really the same. Here τ is the proper time as measured by the infalling object, and r is not the radial distance.

Solving the Equation

$$\boxed{\frac{\mathrm{d}^2 r}{\mathrm{d}\tau^2} = -\frac{GM}{r^2} \quad .}$$

Like Newton's equation, multiply by $\mathrm{d}r/\mathrm{d}\tau$, and it can then be written as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left\{\frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - \frac{GM}{r}\right\} = 0 \ .$$

Quantity in curly brackets is conserved. Initial value (on release from rest at r_0) is $-GM/r_0$, so it always has this value. Then

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = -\sqrt{2GM\left(\frac{1}{r} - \frac{1}{r_0}\right)} = -\sqrt{\frac{2GM(r_0 - r)}{r r_0}} \ .$$

Repeating,

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = -\sqrt{2GM\left(\frac{1}{r} - \frac{1}{r_0}\right)} = -\sqrt{\frac{2GM(r_0 - r)}{rr_0}}\ .$$

Bring all r-dependent factors to one side, and bring $\mathrm{d}\tau$ to the other side, and integrate:

$$\begin{aligned}\tau(r_f) &= -\int_{r_0}^{r_f} \mathrm{d}r \sqrt{\frac{rr_0}{2GM(r_0 - r)}} \\ &= \sqrt{\frac{r_0}{2GM}}\left\{r_0 \tan^{-1}\left(\sqrt{\frac{r_0 - r_f}{r_f}}\right) + \sqrt{r_f(r_0 - r_f)}\right\}\ ,\end{aligned}$$

where $\tan^{-1} \equiv \arctan$.

Conclusion: object will reach $r = 0$ in a finite proper time τ.

$$\tau(r_f) = \sqrt{\frac{r_0}{2GM}}\left\{r_0 \tan^{-1}\left(\sqrt{\frac{r_0 - r_f}{r_f}}\right) + \sqrt{r_f(r_0 - r_f)}\right\} \; .$$

Setting $r_f = 0$ to find the proper time when the object reaches $r = 0$,

$$\tau(0) = \sqrt{\frac{r_0}{2GM}}\left\{r_0 \tan^{-1}(\infty) + 0\right\}$$

$$= \boxed{\frac{\pi}{2}\sqrt{\frac{r_0^3}{2GM}} \; .}$$

Falling from the Schwarzschild Horizon to $r = 0$

Recall,

$$\boxed{\tau(0) = \frac{\pi}{2}\sqrt{\frac{r_0^3}{2GM}}\ .}$$

For $r_0 = R_S$,

$$\boxed{\tau = \frac{\pi GM}{c^3}\ .}$$

For $r_0 = R_S$,

$$\boxed{\tau = \frac{\pi G M}{c^3}\ .}$$

For the Sun, this gives

$$\boxed{\tau = 1.55 \times 10^{-5}\ \text{s.}}$$

For the black hole in the center of our galaxy,

$$\boxed{\tau = 6.34\ \text{s.}}$$

Note that inside the black hole,

$$
\boxed{\begin{aligned} ds^2 = -c^2 d\tau^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 \\ + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2) \ , \end{aligned}}
$$

but

$$
\boxed{\left(1 - \frac{2GM}{rc^2}\right) < 0 \ ,}
$$

which implies that t is spacelike, and r is timelike! The calculation that we just did is still correct. The singularity at $r = 0$ cannot be avoided for the same reason that we cannot prevent ourselves from reaching tomorrow!

But Coordinate Time t is Different!

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = -\sqrt{2GM\left(\frac{1}{r} - \frac{1}{r_0}\right)} = -\sqrt{\frac{2GM(r_0 - r)}{r r_0}} \ .$$

$$c^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2 \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 \ .$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t} = \frac{\mathrm{d}r/\mathrm{d}\tau}{\mathrm{d}t/\mathrm{d}\tau}$$

$$= \frac{\mathrm{d}r/\mathrm{d}\tau}{\sqrt{h^{-1}(r) + c^{-2}h^{-2}(r)\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2}} \ ,$$

where $h^{-1}(r) \equiv 1/h(r)$, not the inverse function, and

$$h(r) \equiv 1 - \frac{R_S}{r} = 1 - \frac{2GM}{rc^2} \ .$$

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = -\sqrt{2GM\left(\frac{1}{r} - \frac{1}{r_0}\right)} = -\sqrt{\frac{2GM(r_0 - r)}{r r_0}} \ .$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r/\mathrm{d}\tau}{\sqrt{h^{-1}(r) + c^{-2}h^{-2}(r)\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2}} \ ,$$

where

$$h(r) \equiv 1 - \frac{R_S}{r} = 1 - \frac{2GM}{rc^2} \ .$$

Look at behavior near horizon; $h^{-1}(r)$ blows up:

$$h^{-1}(r) = \frac{r}{r - R_S} \approx \frac{R_S}{r - R_S} \ .$$

Denominator of $\mathrm{d}r/\mathrm{d}t$ is dominated by 2nd term, which gives

$$\frac{\mathrm{d}r}{\mathrm{d}t} \approx -ch(r) = -c\left(\frac{r - R_S}{R_S}\right) \ .$$

Repeating,

$$\frac{\mathrm{d}r}{\mathrm{d}t} \approx -c\left(\frac{r - R_S}{R_S}\right) \; .$$

Rearranging,

$$\mathrm{d}t = -\frac{R_S}{c}\frac{\mathrm{d}r}{r - R_S} \; .$$

We can find the time needed to fall from some r_i near the horizon, to a smaller r_f which is nearer to the horizon:

$$t(r_f) \approx -\frac{R_S}{c}\int_{r_i}^{r_f} \frac{\mathrm{d}r'}{r' - R_S} \approx \boxed{\frac{R_S}{c}\ln\left(\frac{r_i - R_S}{r_f - R_S}\right) \quad .}$$

Thus t diverges logarithmically as $r_f \to R_S$, so the object does not reach R_S for any finite value of t.

8.286 Class 14
October 26, 2022

BLACK-BODY RADIATION
AND
THE EARLY HISTORY OF THE UNIVERSE

$$\boxed{E = mc^2}$$

$$\boxed{E = mc^2}$$

★ THE most famous equation in physics.

$$\boxed{E = mc^2}$$

- THE most famous equation in physics. But I was not able to find any actual surveys.

$$\boxed{E = mc^2}$$

- ☆ THE most famous equation in physics. But I was not able to find any actual surveys.

- ☆ **Meaning: Mass and energy are equivalent.** They are just two different ways of expressing exactly the same thing. The total energy of any system is equal to the total mass of the system — sometimes called the relativistic mass — times c^2, the square of the speed of light.

$$\boxed{E = mc^2}$$

- THE most famous equation in physics. But I was not able to find any actual surveys.
- **Meaning: Mass and energy are equivalent.** They are just two different ways of expressing exactly the same thing. The total energy of any system is equal to the total mass of the system — sometimes called the relativistic mass — times c^2, the square of the speed of light.
- One can imagine measuring the mass/energy of an object in either kilograms, joules, or kilowatt-hours, with

$$\boxed{1 \text{ kg} = 8.9876 \times 10^{16} \text{ joule} = 2.497 \times 10^{10} \text{ kW-hr.}}$$

$E = mc^2$ and the World Power Supply

- The total amount of power produced in the world, on average, is about 1.89×10^{10} kW, according to the International Energy Agency (2020).
- This amounts to about 2.5 kW per person.
- If a 15 gallon tank of gasoline could be converted *entirely* into usable energy, it would power the world for $2\frac{1}{2}$ days.
- However, it is not so easy! Even with nuclear power, when a uranium-235 nucleus undergoes fission, only about 0.09% of its mass is converted to energy.

$E = mc^2$ and Particle Masses

☆ Nuclear and particle physicists tend to measure the mass of elementary particles in energy units, usually using either MeV (10^6 eV) or GeV (10^9 eV) as the unit of energy, where

$$1\,\mathrm{eV} = 1 \text{ electron volt} = 1.6022 \times 10^{-19}\ \mathrm{J},$$

and then

$$1\,\mathrm{GeV} = 1.7827 \times 10^{-27}\ \mathrm{kg} \cdot c^2\ .$$

The mass of a proton is 0.938 GeV$/c^2$, and the mass of an electron is 0.511 MeV$/c^2$.

Energy and Momentum in Special Relativity

- ☆ We have talked about the kinematic consequences of special relativity (time dilation, Lorentz contraction, and the relativity of simultaneity), but now we need to bring in the dynamical consequences, involving energy and momentum.
- ☆ In special relativity, the definitions of energy and momentum are different from those in Newtonian mechanics.
- ☆ Why? Because special relativity is based on the principle that the laws of physics in any inertial reference frame are the same, and furthermore, in order for the speed of light be the same in any inertial reference frame, these frames cannot be related to each other as in Newtonian physics. They must instead be related by Lorentz transformations, which take into account the kinematic effects mentioned above.

- Two important laws of physics are the conservation of energy and momentum.
- If energy and momentum kept their Newtonian definitions, then, if they were conserved in one frame, they would not be conserved in other frames.
- The requirement that the conservation equations hold in all frames requires the standard special relativity definitions.

Energy, Momentum, and the Energy-Momentum Four-Vector

- The energy-momentum four-vector is defined by starting with the momentum three-vector $(p^1, p^2, p^3) \equiv (p^x, p^y, p^z)$, and appending a fourth component

$$p^0 = \frac{E}{c} \ ,$$

so the four-vector can be written as

$$\boxed{p^\mu = \left(\frac{E}{c}, \vec{p}\right) \ .}$$

As with the three-vector momentum, the energy-momentum four-vector can be defined for a system of particles as the sum of the vectors for the individual particles.

- ★ The 4-vector p^μ transforms, when we change frames of reference, according to the Lorentz transformation, exactly like the 4-vector $x^\mu = (ct, \vec{x})$.
- ★ Furthermore, the total energy-momentum 4-vector is conserved — in any inertial frame of reference.

Relation of Energy and Momentum to Rest Mass and Velocity

The mass of a particle in its own rest frame is called its rest mass, which we denote by m_0. At velocity $\vec{v}$

$$\vec{p} = \gamma m_0 \vec{v} \; ,$$

$$E = \gamma m_0 c^2 = \sqrt{(m_0 c^2)^2 + |\vec{p}|^2 c^2} \; ,$$

where as usual γ is defined by

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \; .$$

Lorentz Invariance of p^2

$$\vec{p} = \gamma m_0 \vec{v} \ ,$$

$$E = \gamma m_0 c^2 = \sqrt{\left(m_0 c^2\right)^2 + \left|\vec{p}\right|^2 c^2} \ ,$$

Like the Lorentz-invariant interval that we discussed as $\mathrm{d}s^2 = -c^2 \mathrm{d}t^2 + \mathrm{d}\vec{x}^2$, the energy-momentum four-vector has a Lorentz-invariant square:

$$\boxed{p^2 \equiv \left|\vec{p}\right|^2 - \left(p^0\right)^2 = \left|\vec{p}\right|^2 - \frac{E^2}{c^2} = -\left(m_0 c\right)^2 \ .}$$

For a particle at rest,

$$\boxed{E = m_0 c^2 \ .}$$

Energy Exchange in a Simple Chemical Reaction

Consider the reaction

$$\boxed{p + e^- \longrightarrow H + \gamma\ .}$$

Assuming that the proton and electron begin at rest, and ignoring the very small kinetic energy of the hydrogen atom when it recoils from the emitted photon, conservation of energy implies that

$$\boxed{m_H = m_p + m_e - E_\gamma / c^2\ .}$$

The energy given off when the proton and electron bind is called the **binding energy** of the hydrogen atom. It is 13.6 eV.

Relativistic Mass

☆ Since $E = mc^2$, we can define the *relativistic mass* of any particle or system as simply

$$\boxed{m_{\rm rel} \equiv \frac{E}{c^2}\ .}$$

☆ Some authors avoid using the concept of relativistic mass, reserving the word "mass" to mean rest mass m_0. Relativistic mass is certainly a redundant concept, since anything that can be described in terms of $m_{\rm rel}$ can also be described in terms of E.

☆ For cosmology the concept of relativistic mass will be helpful, since relativistic mass is the source of gravity. By calling E/c^2 a mass, we are indicating our recognition that it is the source of gravity.

The Source of Gravity in General Relativity

This is beyond the level of what we need, but for those who are interested, I mention that the Einstein field equations imply that the source of gravitational fields is the *energy-momentum tensor* $T^{\mu\nu}$, where μ and ν are 4-vector indices that take on values from 0 to 3.

$T^{00} = u =$ energy density,

$T^{0i} = T^{i0}$ is $\frac{1}{c}$ times the flow of energy in the i'th direction (i=1,2,3) and is also c times the density of the i'th component of momentum,

$T^{ij} = T^{ji}$ is the flow in the j'th direction of the i'th component of momentum. T^{ij} is often diagonal, with $T^{ij} = p\,\delta^{ij}$, where p is the pressure.

For a homogeneous, isotropic universe model, only u and p will serve as sources for gravity.

Mass of Radiation

☆ Electromagnetic radiation has energy. The energy density is given by

$$u = \frac{1}{2}\left[\epsilon_0 \left|\vec{E}\right|^2 + \frac{1}{\mu_0}\left|\vec{B}\right|^2\right] \; .$$

We won't need this equation, but we need to know that electromagnetic radiation **has** an energy density u.

☆ Energy density implies a (relativistic) mass density

$$\rho = u/c^2 \; .$$

(*Relativistic mass* is defined to be the energy divided by c^2.)

Energy and Momentum of Photons

Photons have zero rest mass.

In general,

$$\boxed{p^2 = |\vec{p}\,|^2 - \frac{E^2}{c^2} = -\,(m_0 c)^2 \;\;,}$$

but for photons, $m_0 = 0$, so

$$\boxed{|\vec{p}\,|^2 - \frac{E^2}{c^2} = 0 \;, \quad \text{or} \quad E = c|\vec{p}\,| \;.}$$

Radiation in an Expanding Universe

- From the end of inflation (maybe about 10^{-35} second, to be discussed later) until stars form, the number of photons is almost exactly conserved.

- Therefore,

$$\boxed{n_\gamma \propto \frac{1}{a^3(t)}\ .}$$

- But the frequency of each photon redshifts:

$$\boxed{\nu \propto \frac{1}{a(t)}\ .}$$

$$\boxed{n_\gamma \propto \frac{1}{a^3(t)} \; , \qquad \nu \propto \frac{1}{a(t)} \; .}$$

☆ But according to quantum mechanics, the energy of each photon is

$$\boxed{E = h\nu \; ,}$$

so the energy of each photon is proportional to $1/a(t)$.

☆ Finally,

$$n_\gamma \propto \frac{1}{a^3(t)} \;\; , \quad E_\gamma \propto \frac{1}{a(t)} \quad \Longrightarrow \quad \boxed{\rho_\gamma = \frac{u_\gamma}{c^2} \propto \frac{1}{a^4(t)} \; .}$$

The Radiation Dominated Era

Radiation energy density today (including photons and neutrinos):

$$\boxed{u_r = 7.01 \times 10^{-14}\ \mathrm{J/m^3}\ , \qquad \rho_r = u_r/c^2 = 7.80 \times 10^{-34}\ \mathrm{g/cm^3}\ .}$$

Total mass density today, ρ_0, is equal to within uncertainties to the critical density,

$$\boxed{\rho_c = \frac{3H_0^2}{8\pi G} = 1.88\, h_0^2\ \times 10^{-29}\ \mathrm{g/cm^3}\ ,}$$

where

$$\boxed{H_0 = 100\, h_0\, \mathrm{km\text{-}s^{-1}\text{-}Mpc^{-1}}\ , \qquad h_0 \approx 0.67}\ ,$$

which gives the present value of Ω_r as $\boxed{\Omega_r \approx 9.2 \times 10^{-5}}$.

Since $\rho_r \propto 1/a^4(t)$, while $\rho_m \propto 1/a^3(t)$.

ρ_m = mass density of nonrelativistic matter, baryonic matter plus dark matter. It follows that

$$\boxed{\rho_r/\rho_m \propto 1/a(t)\ .}$$

Today $\rho_m \approx 0.30\rho_c$, so $\rho_r/\rho_m \approx 9.2 \times 10^{-5}/0.30 \approx 3.1 \times 10^{-4}$. Thus

$$\frac{\rho_r(t)}{\rho_m(t)} = \frac{a(t_0)}{a(t)} \times 3.1 \times 10^{-4}\ .$$

$t_{\rm eq}$ is defined to be the time of matter-radiation equality. Thus

$$\frac{\rho_r(t_{\rm eq})}{\rho_m(t_{\rm eq})} \equiv 1 = \frac{a(t_0)}{a(t_{\rm eq})} \times 3.1 \times 10^{-4}\ .$$

Since $a(t_0)/a(t_{\rm eq}) = 1 + z_{\rm eq}$,

$$\boxed{z_{\rm eq} = \frac{1}{3.1 \times 10^{-4}} - 1 \approx 3200\ .}$$

Since $a(t_0)/a(t_{\rm eq}) = 1 + z_{\rm eq}$,

$$\boxed{z_{\rm eq} = \frac{1}{3.1 \times 10^{-4}} - 1 \approx 3200\ .}$$

Time of matter-radiation equality:

We are not ready to calculate this accurately, but for now we can estimate it by assuming that between $t_{\rm eq}$ and now, $a(t) \propto t^{2/3}$, as in a matter-dominated flat universe. Then

$$(t_{\rm eq}/t_0)^{2/3} = 3.1 \times 10^{-4}\ ,$$

so

$$\boxed{t_{\rm eq} = 5.5 \times 10^{-6}\, t_0 = 5.5 \times 10^{-6} \times 13.8 \text{ Gyr} \approx 75{,}000 \text{ years.}}$$

Ryden (p. 96) gives 50,000 years, which is more accurate.

Dynamics of the Radiation-Dominated Era

$$\rho \propto \frac{1}{a^3} \implies \dot\rho = -3\frac{\dot a}{a}\rho\ , \quad \rho(t) \propto \frac{1}{a^4(t)} \implies \dot\rho = -4\frac{\dot a}{a}\rho\ .$$

$\dot\rho$ and pressure p: (Problem 1, Problem Set 7)

$$dU = -p\,dV \implies \frac{dU}{dt} = -p\frac{dV}{dt}$$

$$= \frac{d}{dt}\left(a^3\rho c^2\right) = -p\frac{d}{dt}(a^3)$$

$$\implies \dot\rho = -3\frac{\dot a}{a}\left(\rho + \frac{p}{c^2}\right)\ .$$

Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}$$

$$\ddot{a} = -\frac{4\pi}{3}G\rho a \ ,$$

$$\dot{\rho} = -3\frac{\dot{a}}{a}\,\rho$$

$$\begin{pmatrix}\text{matter-dominated} \\ \text{universe}\end{pmatrix}$$

Any two of the above equations implies the third. So they become inconsistent if

$$\boxed{\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \ .}$$

$$\left(\frac{\dot a}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}\ , \qquad \ddot a = -\frac{4\pi}{3}G\rho a\ .$$

Any two of the above equations implies the third. So they become inconsistent if

$$\dot\rho = -3\frac{\dot a}{a}\left(\rho + \frac{p}{c^2}\right)\ .$$

So, if we believe the equation for $\dot\rho$, we must modify one of the two Friedmann equations. First order equation represents conservation of energy: pressure does not belong! (Pressures can change suddenly, as when dynamite explodes, so it does not make sense to have pressure in a conservation equation.) So modify the 2nd order equation, deriving it from the first order equation and the $\dot\rho$ equation:

$$\ddot a = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a\ .$$

8.286 Class 15
October 31, 2022

BLACK-BODY RADIATION
AND
THE EARLY HISTORY OF THE UNIVERSE, PART 2

PROBLEM 3: THE CRUNCH OF A CLOSED, MATTER-DOMINATED UNIVERSE *(25 points)*

This is Problem 5.7 (Problem 6.5 in the first edition) from Barbara Ryden's ***Introduction to Cosmology****, with some paraphrasing to make it consistent with the language used in lecture.*

Consider a closed universe containing only nonrelativistic matter. This is the closed universe discussed in Lecture Notes 4, and it is also the "Big Crunch" model discussed in Ryden's section Section 5.4.1 (Section 6.1 in the first edition). At some time during the contracting phase (i.e., when $\theta > \pi$), an astronomer named Elbbuh Niwde discovers that nearby galaxies have blueshifts ($-1 \le z < 0$) proportional to their distance. He then measures the present values of the Hubble expansion rate, H_0, and the mass density parameter, Ω_0. He finds, of course, that $H_0 < 0$ (because he is in the contracting phase) and $\Omega_0 > 1$ (because the universe is closed). **In terms of H_0 and Ω_0, how long a time will elapse between Dr. Niwde's observation at $t = t_0$ and the final Big Crunch at $t = t_{\text{Crunch}} = 2\pi\alpha/c$?** Assuming that Dr. Niwde is able to observe all objects within his horizon, **what is the most blueshifted (i.e., most negative) value of z that Dr. Niwde is able to see? What is the lookback time to an object with this blueshift?** (By lookback time, one means the difference between the time of observation t_0 and the time at which the light was emitted.)

Radiation In An Expanding Universe

☆ As the universe expands,

$$\boxed{n_\gamma \propto \frac{1}{a^3(t)}\ , \quad \nu \propto \frac{1}{a(t)}\ , \quad E_\gamma = h\nu \quad \Longrightarrow \quad u_\gamma \propto \rho_\gamma \propto \frac{1}{a^4(t)}\ ,}$$

where n_γ = number density of photons, ν = frequency of any one photon, E_γ = energy of any one photon, and u_γ and ρ_γ are the energy density and mass density of photons, respectively.

Review from the previous lecture

The Radiation Dominated Era

Radiation energy density today (including photons and neutrinos):

$$\boxed{u_r = 7.01 \times 10^{-14}\ \mathrm{J/m^3}\ , \qquad \rho_r = u_r/c^2 = 7.80 \times 10^{-34}\ \mathrm{g/cm^3}\ .}$$

Total mass density today, ρ_0, is equal, to within uncertainties of a fraction of a percent, to the critical density,

$$\boxed{\rho_c = \frac{3H_0^2}{8\pi G} = 1.88\, h_0^2 \times 10^{-29}\ \mathrm{g/cm^3}\ ,}$$

where

$$\boxed{H_0 = 100\, h_0\, \mathrm{km\text{-}s^{-1}\text{-}Mpc^{-1}}\ , \qquad h_0 \approx 0.67}\ ,$$

which gives the present value of Ω_r as $\boxed{\Omega_r \approx 9.2 \times 10^{-5}}$.

Since $\rho_r \propto 1/a^4(t)$, while $\rho_m \propto 1/a^3(t)$.

ρ_m = mass density of nonrelativistic matter: baryonic matter plus dark matter. It follows that

$$\boxed{\rho_r/\rho_m \propto 1/a(t)\ .}$$

Today $\rho_m \approx 0.30\rho_c$.

Carrying out the calculations, we found that matter-radiation equality occurred at

$$\boxed{z_{\rm eq} \approx 3200\ ,}$$

and

$$\boxed{t_{\rm eq} = 5.5 \times 10^{-6}\, t_0 = 5.5 \times 10^{-6} \times 13.8 \text{ Gyr} \approx 75,000 \text{ years.}}$$

Ryden (p. 96) gives 50,000 years, which is more accurate. We have not yet included a cosmological constant, and we treated the universe as completely matter-dominated even just after $t_{\rm eq}$.

Dynamics of the Radiation-Dominated Era

$$\rho \propto \frac{1}{a^3} \implies \dot\rho = -3\frac{\dot a}{a}\rho \ , \quad \rho(t) \propto \frac{1}{a^4(t)} \implies \dot\rho = -4\frac{\dot a}{a}\rho \ .$$

$\dot\rho$ and pressure p: (Problem 1, Problem Set 7)

$$dU = -p\,dV \quad \implies \quad \frac{dU}{dt} = -p\frac{dV}{dt}$$

$$= \frac{d}{dt}\left(a^3\rho c^2\right) = -p\frac{d}{dt}(a^3)$$

$$\implies \quad \dot\rho = -3\frac{\dot a}{a}\left(\rho + \frac{p}{c^2}\right) \ .$$

Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}$$

$$\ddot{a} = -\frac{4\pi}{3}G\rho a \ ,$$

$$\dot{\rho} = -3\frac{\dot{a}}{a}\,\rho$$

$\left(\begin{array}{c}\text{matter-dominated}\\ \text{universe}\end{array}\right)$

Any two of the above equations implies the third. So they become inconsistent if

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \ .$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}\ , \qquad \ddot{a} = -\frac{4\pi}{3}G\rho a\ , \qquad \dot{\rho} = -3\frac{\dot{a}}{a}\rho\ .$$

Any two of the above equations implies the third. So they become inconsistent if

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right)\ .$$

So, if we believe the equation for $\dot{\rho}$, we must modify one of the two Friedmann equations. First order equation represents conservation of energy: pressure does not belong! (Pressures can change suddenly, as when dynamite explodes, so it does not make sense to have pressure in a conservation equation.) So modify the 2nd order equation, deriving it from the first order equation and the $\dot{\rho}$ equation:

$$\ddot{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a\ .$$

Summary: Complete Friedmann Equations and Energy Conservation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3} G\rho - \frac{kc^2}{a^2}$$

$$\ddot{a} = -\frac{4\pi}{3} G \left(\rho + \frac{3p}{c^2}\right) a$$

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) .$$

The items in red are new.

Dynamics of a Flat Radiation-dominated Universe

$$H^2 = \frac{8\pi G}{3}\rho \ , \ \ \rho \propto 1/a^4 \qquad \Longrightarrow \qquad \left(\frac{\dot{a}}{a}\right)^2 = \frac{\text{const}}{a^4} \ .$$

Then

$$a\,\mathrm{d}a = \sqrt{\text{const}}\,\mathrm{d}t \qquad \Longrightarrow \qquad \frac{1}{2}a^2 = \sqrt{\text{const}}\,t + \text{const}' \ .$$

So, setting our clocks so that $\text{const}' = 0$,

$$\boxed{a(t) \propto \sqrt{t} \qquad \text{(flat radiation-dominated)} \ .}$$

$$\boxed{H(t) = \frac{\dot{a}}{a} = \frac{1}{2t} \qquad \text{(flat radiation-dominated)}\ .}$$

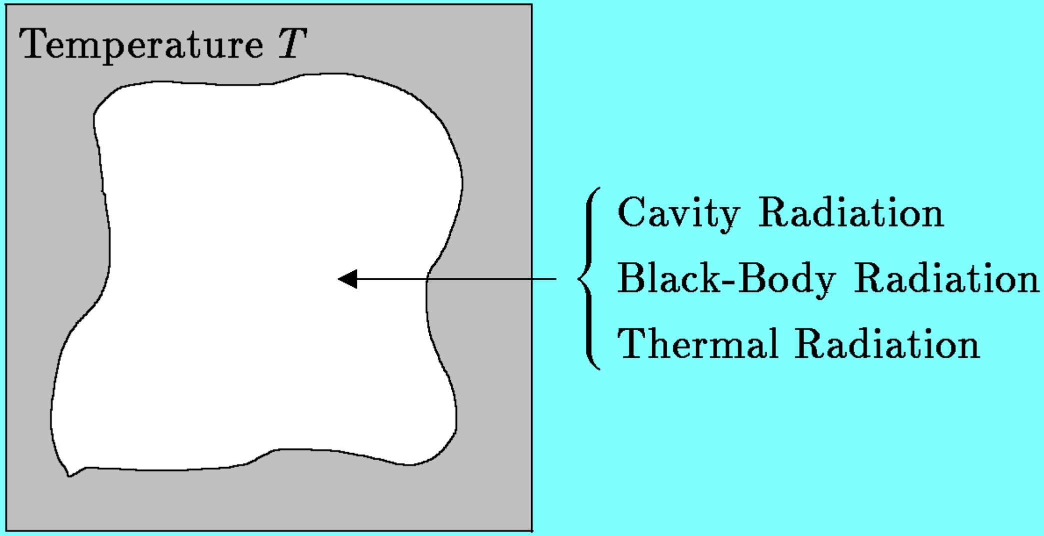
$$\ell_{p,\text{horizon}}(t) = a(t)\int_0^t \frac{c}{a(t')}dt'$$

$$= \boxed{2ct \qquad \text{(flat radiation-dominated)}\ .}$$

$$H^2 = \frac{8\pi G}{3}\rho \quad \Longrightarrow \quad \boxed{\rho = \frac{3}{32\pi G t^2}\ .}$$

Black Body Radiation

- If a cavity is carved out of any material, and the walls are kept at a uniform temperature T, then the cavity will fill with radiation.



- If no radiation can get through the wall, then the energy density and spectrum of the radiation is determined by T alone — the material of the wall is irrelevant.
- The radiation is known as cavity radiation, black-body radiation, or thermal radiation.
- It can be thought of simply as radiation at temperature T.

Why Is It Called Black-Body?

- A black body at temperature T in empty space emits radiation with exactly this intensity and spectrum.
- Definitions:
  - A *black* object absorbs all light that hits it; none is reflected or transmitted.
  - Reflection vs. emission: *reflection* is immediate. If the body absorbs radiation and emits it later, that is *emission*.
- Equilibrium: if a black body were placed in the cavity, it would reach an equilibrium in which no further energy would be exchanged. The body would be at the same temperature T as the box and the cavity radiation.
- Since the black body absorbs all the radiation that hits it, it must emit exactly this much radiation.

- Furthermore, in every frequency interval the block must emit exactly as much radiation as it absorbs.

☆ Furthermore, in every frequency interval the block must emit exactly as much radiation as it absorbs.

Otherwise, we could imagine surrounding the body by a filter that transmits only in this frequency interval, and otherwise reflects. If the emission in this interval did not match the aborption, the body would then become hotter or colder than T, which violates a basic property of thermal equilibrium — once it is reached, the temperature will remain uniform, unless energy is exchanged with some external mechanism.

- ☆ Furthermore, in every frequency interval the block must emit exactly as much radiation as it absorbs.

  Otherwise, we could imagine surrounding the body by a filter that transmits only in this frequency interval, and otherwise reflects. If the emission in this interval did not match the aborption, the body would then become hotter or colder than T, which violates a basic property of thermal equilibrium — once it is reached, the temperature will remain uniform, unless energy is exchanged with some external mechanism.

- ☆ Since the black body reflects nothing, all of the emitted radiation is thermal radiation, which will continue even if the body is taken out of the cavity.

- Furthermore, in every frequency interval the block must emit exactly as much radiation as it absorbs.

  Otherwise, we could imagine surrounding the body by a filter that transmits only in this frequency interval, and otherwise reflects. If the emission in this interval did not match the aborption, the body would then become hotter or colder than T, which violates a basic property of thermal equilibrium — once it is reached, the temperature will remain uniform, unless energy is exchanged with some external mechanism.

- Since the black body reflects nothing, all of the emitted radiation is thermal radiation, which will continue even if the body is taken out of the cavity.

- Thus, a black body at temperature T will emit with exactly the same intensity and spectrum as the radiation in the cavity.

Vague Description of the Black-Body Radiation Calculation

- We will leave the full derivation of black-body radiation to some stat mech class.
- But here we will summarize the basic ideas.
- **Prelude: The "equipartition theorem" of classical stat mech:** each degree of freedom of a system at temperature T acquires a mean thermal energy of $\frac{1}{2}kT$, where $k =$ Boltzmann constant $= 8.617 \times 10^{-5}$ eV/K. For example, a gas of spinless particles has 3 degrees of freedom per atom: the x, y, and z components of velocity. In thermal equilbrium, the thermal energy is $\frac{3}{2}kT$ per particle. A harmonic oscillator has 2 degrees of freedom: its kinetic and potential energies.

★ Equipartition and the electromagnetic radiation:

☆ Equipartition and the electromagnetic radiation:

- Imagine describing the electromagnetic field inside a rectangular box. With reflecting boundary conditions, the field can be described by standing waves, each with an integral number of half wavelengths in each of the 3 directions.

☆ Equipartition and the electromagnetic radiation:

- Imagine describing the electromagnetic field inside a rectangular box. With reflecting boundary conditions, the field can be described by standing waves, each with an integral number of half wavelengths in each of the 3 directions.

- There are 2 polarizations (right and left circular polarization, or x and y linear polarization — these are two different bases for the same space of solutions; any polarization can be written as a superposition of left and right circular polarization, **OR** x and y polarization; either way, it counts as TWO polarizations). Each standing wave, with a specified polarization, is called a *mode*. Each mode is 2 degrees of freedom, like a harmonic oscillator.

☆ **Equipartition and the electromagnetic radiation:**

- Imagine describing the electromagnetic field inside a rectangular box. With reflecting boundary conditions, the field can be described by standing waves, each with an integral number of half wavelengths in each of the 3 directions.

- There are 2 polarizations (right and left circular polarization, or x and y linear polarization — these are two different bases for the same space of solutions; any polarization can be written as a superposition of left and right circular polarization, **OR** x and y polarization; either way, it counts as TWO polarizations). Each standing wave, with a specified polarization, is called a *mode*. Each mode is 2 degrees of freedom, like a harmonic oscillator.

- **Jeans Catastrophe:** The number of modes is **infinite**, since there is no shortest wavelength. If classical physics applied, the electromagnetic field could never reach thermal equilibrium. Instead, it would continue to absorb energy, exciting shorter and shorter wavelength modes. It would be an infinite heat sink, absorbing all thermal energy.

☆ Quantum Theory to the Rescue:

- Classically, each mode can be excited by any amount.
- Quantum mechanically, however, a harmonic oscillator with frequency ν can only acquire energy in lumps of size $h\nu$. For the E&M field, each excitation of energy $h\nu$ is a *photon*.

☆ Quantum Theory to the Rescue:

- Classically, each mode can be excited by any amount.
- Quantum mechanically, however, a harmonic oscillator with frequency ν can only acquire energy in lumps of size $h\nu$. For the E&M field, each excitation of energy $h\nu$ is a *photon*.
- For modes for which $h\nu \ll kT$, the classical physics works, and each mode acquires energy kT. (Note: Lecture Notes 6 incorrectly states that for $h\nu \ll kT$, each mode acquires energy $\frac{1}{2}kT$ — it's really kT, with the 2 degrees of freedom of a harmonic oscillator.)

☆ Quantum Theory to the Rescue:

- Classically, each mode can be excited by any amount.
- Quantum mechanically, however, a harmonic oscillator with frequency ν can only acquire energy in lumps of size $h\nu$. For the E&M field, each excitation of energy $h\nu$ is a *photon*.
- For modes for which $h\nu \ll kT$, the classical physics works, and each mode acquires energy kT. (Note: Lecture Notes 6 incorrectly states that for $h\nu \ll kT$, each mode acquires energy $\frac{1}{2}kT$ — it's really kT, with the 2 degrees of freedom of a harmonic oscillator.)
- For modes with $h\nu \gg kT$, the typical energy available ($\sim kT$) is much smaller than the minimum possible excitation ($h\nu$). These modes are excited only very rarely. The Jeans catastrophe is avoided, and the total energy density is finite.

Black-Body Radiation: Results

Energy Density:

$$\boxed{u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}\ ,}$$

where

$$\boxed{\hbar = \frac{h}{2\pi} = 1.055 \times 10^{-27}\ \text{erg-sec} = 6.582 \times 10^{-16}\ \text{eV-sec}\ ,}$$

and

$$\boxed{g = 2 \quad \text{(for photons)}\ .}$$

$$u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}\ , \qquad g = 2 \quad \text{(for photons)}\ .$$

The factor of g is introduced so that the formula will be reusable. We will soon be talking about thermal radiation of other kinds of particles (neutrinos, e^+e^- pairs, and more!), and we'll be able to use the same formula, with different values of g.

$$u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}\ , \qquad g = 2 \quad \text{(for photons)}\ .$$

The factor of g is introduced so that the formula will be reusable. We will soon be talking about thermal radiation of other kinds of particles (neutrinos, e^+e^- pairs, and more!), and we'll be able to use the same formula, with different values of g. Very ecological.

$$u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}\;, \qquad g = 2 \quad \text{(for photons)}\;.$$

The factor of g is introduced so that the formula will be reusable. We will soon be talking about thermal radiation of other kinds of particles (neutrinos, e^+e^- pairs, and more!), and we'll be able to use the same formula, with different values of g. Very ecological.

For photons, $g = 2$ because the photon has two polarizations, or equivalently, two spin states.

Other Properties

Pressure: $$p = \frac{1}{3} u \ .$$

Number Density: $$n = g^* \frac{\zeta(3)}{\pi^2} \frac{(kT)^3}{(\hbar c)^3} \ ,$$

where $\zeta(3)$ is the Riemann zeta function evaluated at 3,

$$\zeta(3) = \frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots \approx 1.202 \ ,$$

and

$$g^* = 2 \quad \text{(for photons)} \ .$$

Number Density: $n = g^* \dfrac{\zeta(3)}{\pi^2} \dfrac{(kT)^3}{(\hbar c)^3}$,

where $\zeta(3)$ is the Riemann zeta function evaluated at 3,

$$\zeta(3) = \frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots \approx 1.202 \ ,$$

and

$$g^* = 2 \quad \text{(for photons)} \ .$$

---

g^* is used in the equation for the number density, rather than g, again to maximize reusability. For photons, $g^* = g$, but that won't be true for all particles.

ENTROPY!!

- Entropy is often described as a measure of the "disorder" of the state of a physical system. Roughly, the entropy of a system is k times the logarithm of the number of microscopic quantum states that are consistent with its macroscropically observed state.

ENTROPY!!

- ☆ Entropy is often described as a measure of the "disorder" of the state of a physical system. Roughly, the entropy of a system is k times the logarithm of the number of microscopic quantum states that are consistent with its macroscropically observed state.

- ☆ Good news: **we will not really need to know what entropy means!** (We will defer to some stat mech class to make this clear.)

ENTROPY!!

- Entropy is often described as a measure of the "disorder" of the state of a physical system. Roughly, the entropy of a system is k times the logarithm of the number of microscopic quantum states that are consistent with its macroscropically observed state.

- Good news: **we will not really need to know what entropy means!** (We will defer to some stat mech class to make this clear.) However, we will make much use of the fact that, as long as a system remains very close to thermal equilibrium, entropy is conserved. When departures from thermal equilibrium occur, the entropy always increases (a principle called the *second law of thermodynamics*).

ENTROPY!!

- ☆ Entropy is often described as a measure of the "disorder" of the state of a physical system. Roughly, the entropy of a system is k times the logarithm of the number of microscopic quantum states that are consistent with its macroscropically observed state.
- ☆ Good news: **we will not really need to know what entropy means!** (We will defer to some stat mech class to make this clear.) However, we will make much use of the fact that, as long as a system remains very close to thermal equilibrium, entropy is conserved. When departures from thermal equilibrium occur, the entropy always increases (a principle called the *second law of thermodynamics*).
- ☆ In our model of the universe, a huge amount of entropy was produced at the end of the period of inflation (to be discussed later), but the subsequent expansion and cooling of the universe happens at nearly constant entropy. Once stars form, entropy production resumes.

Entropy Density of Black-Body Radiation

The entropy density s of black-body radiation is given by

$$\boxed{s = g\frac{2\pi^2}{45}\,\frac{k^4T^3}{(\hbar c)^3}\ .}$$

The factor of g that appears here is the same g that occurs in the formulas for energy density and pressure. For photons, g is (still) 2.

Note that the entropy density, like the number density, is proportional to T^3. Thus the ratio

$$\boxed{\frac{s}{n} = \frac{g}{g^*}3.60157\,k\ .}$$

For the black-body radiation of photons, entropy is just another way to count photons, with 3.6 k units of entropy per photon.

Neutrinos — A Brief History

- ☆ In 1930, Wolfgang Pauli proposed the existence of the neutrino — an unseen particle that he theorized to explain how beta decay ($n \longrightarrow p + e^-$, inside a nucleus) could be consistent with energy conservation. (Niels Bohr, by contrast, proposed that energy conservation was only valid statistically.) Pauli called it a neutron, while the particle that we know as a neutron was not discovered until 1932, by James Chadwick.
- ☆ In 1934 Enrico Fermi developed a full theory of beta decay, and gave the neutrino its current name ("little neutral one").
- ☆ The neutrino was not seen observationally until 1956 by Clyde Cowan and Frederick Reines at the Savannah River nuclear reactor.
- ☆ Cowan died in 1974 at the age of 54, and Reines was awarded the Nobel Prize for this work in 1995, at the age of 77.

8.286 Class 16
November 2, 2022

BLACK-BODY RADIATION AND THE EARLY HISTORY OF THE UNIVERSE, PART 3

Summary: Complete Friedmann Equations and Energy Conservation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}$$

$$\ddot{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a$$

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) .$$

The items in red are new.

Review from the previous lecture

Black-Body Radiation

$$\text{Energy density:} \quad u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}$$

$$\text{Pressure:} \quad p = \frac{1}{3}u$$

$$\text{Number density:} \quad n = g^*\frac{\zeta(3)}{\pi^2}\,\frac{(kT)^3}{(\hbar c)^3}$$

$$\text{Entropy density:} \quad s = g\frac{2\pi^2}{45}\,\frac{k^4T^3}{(\hbar c)^3}\ ,$$

where for photons

$$\boxed{g = g^* = 2\ ,}$$

But g and g^* will have different values for different particles, to be discussed today.

Neutrinos — A Brief History

- ☆ In 1930, Wolfgang Pauli proposed the existence of the neutrino — an unseen particle that he theorized to explain how beta decay ($n \longrightarrow p + e^-$, inside a nucleus) could be consistent with energy conservation. (Niels Bohr, by contrast, proposed that energy conservation was only valid statistically.) Pauli called it a neutron, while the particle that we know as a neutron was not discovered until 1932, by James Chadwick.
- ☆ In 1934 Enrico Fermi developed a full theory of beta decay, and gave the neutrino its current name ("little neutral one").
- ☆ The neutrino was not seen observationally until 1956 by Clyde Cowan and Frederick Reines at the Savannah River nuclear reactor.
- ☆ Cowan died in 1974 at the age of 54, and Reines was awarded the Nobel Prize for this work in 1995, at the age of 77.

Neutrino Mass, Take 1

- During the 20th century, neutrinos were thought to be massless (rest mass = 0). We now know that they have a very small but nonzero mass, but for the period that we will be discussing now (the radiation-dominated era), the masses are negligible. As long as $mc^2 \ll kT$, the particle will act as if it is massless.
- So, for now (Take 1), we will pretend neutrinos are massless.

Photons are Bosons, Neutrinos are Fermions

- All particles can be divided into these two classes.
- For bosons, any number of particles can exist in the same quantum state. This is what allows photons to build up a classical electromagnetic field, which involves a very large number of photons. A laser in particular concentrates a huge number of photons in a single quantum state.
- For fermions, by contrast, there can be no more than one particle in a given quantum state. Electrons are also fermions — the one-electron-per-quantum-state rule is called the *Pauli Exclusion Principle*, and is responsible for essentially all of chemistry.
- In relativistic quantum field theory, one can prove the *spin-statistics theorem*: all particles with integer spin (in units of $\hbar$) are bosons, and all particles with half-integer spin ($\frac{1}{2}$, $\frac{3}{2}$, etc.) are fermions. (And those are the only possibilities.)

Consequences of Fermi Statistics

Reminder:

$$u = g\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}\ , \qquad n = g^*\frac{\zeta(3)}{\pi^2}\,\frac{(kT)^3}{(\hbar c)^3}\ .$$

- ☆ Because there are fewer states that fermions can occupy, the number density, energy density, pressure, and entropy density for fermions are all reduced.
- ☆ For fermions,

g is reduced by a factor of $7/8$.

g^* is reduced by a factor of $3/4$.

Neutrino Flavors

Neutrinos come in 3 different species, or *flavors*:

$$\begin{aligned}
\text{Electron neutrino } \nu_e: &\qquad e^- + p \longrightarrow n + \nu_e \\
\text{Muon neutrino } \nu_\mu: &\qquad \mu^- + p \longrightarrow n + \nu_\mu \\
\text{Tau neutrino } \nu_\tau: &\qquad \tau^- + p \longrightarrow n + \nu_\tau \ .
\end{aligned}$$

A muon is essentially a heavy electron, with $m_\mu c^2 = 105.7$ MeV, compared to $m_e c^2 = 0.511$ MeV. A tau is a still heavier version of the electron, with $m_\tau c^2 = 1776.9$ MeV.

Neutrino States

- 3 flavors implies a factor of 3 in g and g^*.
- Neutrinos exist as particles and antiparticles, unlike photons, which are their own antiparticles. The particle/antiparticle option leads to a factor of 2 in g and g^*
- While photons can be left or right circularly polarized, neutrinos are always seen to be *left-handed*: the spin is opposite the direction of the momentum. Antineutrinos are always right-handed.

An Aside on Discrete Symmetries

☆ Before the left-handed property of neutrinos was discovered, it was thought that that the laws of physics were invariant under *parity transformations* $(x \to -x, y \to -y, z \to -z)$. But the parity transform of a left-handed neutrino would be a right-handed neutrino, which has never been seen, so the laws of physics are **NOT** parity-invariant.

☆ The handedness of neutrinos is consistent with CP symmetry, charge conjugation time parity. The CP transform of a left-handed neutrino is a right-handed antineutrino — both exist and, as far as we know, behave identically. However, CP symmetry is known to be violated by neutral kaons.

☆ However, CPT symmetry — charge conjugation times parity times time-reversal — is required by relativistic quantum field theory and is believed to be a symmetry of nature.

g and g^* for Neutrinos

$$g_\nu = \underbrace{\frac{7}{8}}_{\text{Fermion factor}} \times \underbrace{3}_{\substack{\text{3 species}\\ \nu_e,\nu_\mu,\nu_\tau}} \times \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{1}_{\text{Spin states}} = \frac{21}{4}\ .$$

$$g^*_\nu = \underbrace{\frac{3}{4}}_{\text{Fermion factor}} \times \underbrace{3}_{\substack{\text{3 species}\\ \nu_e,\nu_\mu,\nu_\tau}} \times \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{1}_{\text{Spin states}} = \frac{9}{2}\ .$$

Hotter Still

If we follow the universe further back in time, we will find that at some point kT becomes large compared to $m_e c^2 = 0.511$ MeV, the rest energy of an electron. Then electron-positron pairs start to behave as massless particles, and contribute to the black-body radiation.

$$g_{e^+e^-} = \underbrace{\frac{7}{8}}_{\text{Fermion factor}} \times \underbrace{1}_{\text{Species}} \times \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{2}_{\text{Spin states}} = \frac{7}{2}\,.$$

$$g^*_{e^+e^-} = \underbrace{\frac{3}{4}}_{\text{Fermion factor}} \times \underbrace{1}_{\text{Species}} \times \underbrace{2}_{\text{Particle/antiparticle}} \times \underbrace{2}_{\text{Spin states}} = 3\,.$$

For 0.511 MeV $\ll kT \ll$ 106 MeV

For electrons, $m_e c^2 = 0.511$ MeV.

For muons, $m_\mu c^2 = 106$ MeV.

For 0.511 MeV $\ll kT \ll$ 106 MeV, electrons and positrons act like massless particles, and only a negligible number of muons would be produced.

The energy density can therefore be calculated from

$$\boxed{g_{\text{tot}} = \underbrace{2}_{\text{photons}} + \underbrace{\frac{21}{4}}_{\text{neutrinos}} + \underbrace{\frac{7}{2}}_{e^+e^-} = 10\,\frac{3}{4}}$$

Energy Density of Radiation Today

★ Temperature of the cosmic microwave background (CMB) today: $\boxed{T_\gamma = 2.7255 \pm 0.0006\ \text{K.}^*}$ This gives $kT_\gamma = 2.35 \times 10^{-4}$ eV.

*D.J. Fixsen, Ap. J. **707,** 916 (2009). Based mainly on the COBE (Cosmic Background Explorer) data, 1989 – 1993.

Energy Density of Radiation Today

- Temperature of the cosmic microwave background (CMB) today: $\boxed{T_\gamma = 2.7255 \pm 0.0006 \text{ K.}^*}$ This gives $kT_\gamma = 2.35 \times 10^{-4}$ eV.
- Continuing our "Take 1" pretense that neutrinos are massless, the radiation that exists in the universe today includes photons and neutrinos.

*D.J. Fixsen, Ap. J. **707,** 916 (2009). Based mainly on the COBE (Cosmic Background Explorer) data, 1989 – 1993.

Energy Density of Radiation Today

- Temperature of the cosmic microwave background (CMB) today: $\boxed{T_\gamma = 2.7255 \pm 0.0006\ \mathrm{K.}^*}$ This gives $kT_\gamma = 2.35 \times 10^{-4}$ eV.

- Continuing our "Take 1" pretense that neutrinos are massless, the radiation that exists in the universe today includes photons and neutrinos.
  But $T_\nu \neq T_\gamma$.

*D.J. Fixsen, Ap. J. **707,** 916 (2009). Based mainly on the COBE (Cosmic Background Explorer) data, 1989 – 1993.

Energy Density of Radiation Today

- Temperature of the cosmic microwave background (CMB) today: $\boxed{T_\gamma = 2.7255 \pm 0.0006 \text{ K.}^*}$ This gives $kT_\gamma = 2.35 \times 10^{-4}$ eV.

- Continuing our "Take 1" pretense that neutrinos are massless, the radiation that exists in the universe today includes photons and neutrinos. **But $T_\nu \neq T_\gamma$.**

- The complication occurs when the e^+e^- pairs "freeze out," (i.e., disappear), as kT falls below 0.511 MeV. This happens around $t = 1$ second. Neutrino interactions become weaker as the temperature falls, and by this time they have become so weak that the neutrinos absorb only a negligible amount of the e^+e^- energy. It essentially all goes into heating the photons, which then become hotter than the neutrinos.

*D.J. Fixsen, Ap. J. **707**, 916 (2009). Based mainly on the COBE (Cosmic Background Explorer) data, 1989 – 1993.

☆ The complication occurs when the e^+e^- pairs "freeze out," (i.e., disappear), as kT falls below 0.511 MeV. This happens around $t = 1$ second. Neutrino interactions become weaker as the temperature falls, and by this time they have become so weak that the neutrinos absorb only a negligible amount of the e^+e^- energy. It essentially all goes into heating the photons, which then become hotter than the neutrinos.

☆ You will calculate this on Problem Set 7. The key is to use *entropy*, not energy, since entropy is simply conserved. Energy density, by contrast, obeys

$$\boxed{\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \ ,}$$

so one needs to calculate the pressure p as the e^+e^- pairs freeze out. That's complicated.

- The result (that you will find) is

$$\boxed{T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma \ .}$$

- This ratio is maintained to the present day, so the total radiation energy density today is

$$\boxed{\begin{aligned} u_{\rm rad,0} &= \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right]\frac{\pi^2}{30}\frac{(kT_\gamma)^4}{(\hbar c)^3} \\ &= 7.01 \times 10^{-14}\ \mathrm{J/m^3}\ , \end{aligned}}$$

which is what we used when we estimated $t_{\rm eq}$, the time of matter-radiation equality.

- We (crudely) found $\sim 75,000$ years. Ryden gives 47,000 years. The Particle Data Group (2020) gives $51,100 \pm 800$ years.

The Real Story of Neutrino Masses

☆ We have not yet measured the mass of a neutrino, but we have seen neutrinos “oscillate” from one flavor to another:

- Electron neutrinos from the Sun arrive at Earth as a mixture of all three flavors.
- Neutrinos produced by cosmic rays in the upper atmosphere have been found to undergo oscillations on their way to ground level.
- Neutrinos produced by reactors and accelerators have been seen to oscillate.

☆ Oscillations require a nonzero mass: essentially because a massless particle experiences an infinite time dilation, so time stops.

☆ The oscillations measure the differences of the squares of the masses.

Neutrino Masses and Quantum Superpositions

- Quantum theory allows for states that are superpositions of other states.
- Neutrinos are produced in states of definite flavor, called ν_e, ν_μ, and ν_τ. (I.e., $e^- + p \longrightarrow n + \nu_e$, while $\mu^- + p$ leads to ν_μ, and $\tau^- + p$ leads to ν_τ.) But these are not states of definite mass!
- The states of definite mass are called ν_1, ν_2, and ν_3.
- Each flavor state is a superposition of all three states of definite mass, and each state of definite mass is a superposition of all three flavor states.

Differences of Squares of Neutrino Masses

As of 2020, the Particle Data Group reports:

$$\Delta m^2_{21}\, c^4 = (7.53 \pm 0.18) \times 10^{-5}\ \text{eV}^2\ ,$$

$$\Delta m^2_{32}\, c^4 = \left(2.546^{+0.034}_{-0.040}\right) \times 10^{-3}\,\text{eV}^2\ ,$$

or

$$\Delta m^2_{32}\, c^4 = (2.453 \pm 0.034) \times 10^{-3}\,\text{eV}^2\ ,$$

where the two options for Δm^2_{32} depend on assumptions about the ordering of the masses. Note that $\sqrt{\Delta m^2_{21}\, c^4} = 8.68 \times 10^{-3}$ eV, and $\sqrt{\Delta m^2_{32}\, c^4} = 0.0505$ eV or 0.0495 eV. Recall that today $kT_\gamma = 2.35 \times 10^{-4}$ eV, which is much smaller.

Does Neutrino Mass Affect Our Calculation of $t_{\rm eq}$?

No!

We wrote

$$\boxed{\begin{aligned} u_{\rm rad,0} &= \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right]\frac{\pi^2}{30}\frac{(kT_\gamma)^4}{(\hbar c)^3} \\ &= 7.01 \times 10^{-14}\ \mathrm{J/m^3}\ , \end{aligned}}$$

but what we really used was

$$\boxed{u_{\rm rad}(t) = \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right]\frac{\pi^2}{30}\frac{(kT_\gamma)^4}{(\hbar c)^3}\left(\frac{a(t_0)}{a(t)}\right)^4\ ,}$$

which is valid for t anywhere near the time $t_{\rm eq}$.

Cosmological Bound on the Sum of ν Masses

☆ From cosmology of large-scale structure, we know that*

$$\boxed{(m_1 + m_2 + m_3)c^2 \leq 0.17 \text{ eV}.}$$

☆ Why? Because neutrinos "free-stream" easily from one place to another. If they carried too much mass, they would even out the mass density and suppress large-scale structure.

*S. R. Choudhury and S. Hannestad, JCAP 2020, No. 7, 037 (2020), arXiv:1907.12598.

Neutrino Mass and Spin States

- The measurements of the mass differences imply that at least 2 of the 3 neutrino masses must be nonzero.
- If the mass of a neutrino is nonzero, then it **cannot** always be left-handed.
- To see this, consider a left-handed neutrino moving in the z direction, with spin in the $-z$ direction. With $m > 0$, it must move slower than c. So an observer can move along the z-axis faster than the neutrino. To such an observer, the momentum of the ν will be in the $-z$ direction, the spin will be in the $-z$ direction, and the ν will appear right-handed.
- How could this right-handed neutrino fit into our theory?

Majorana and Dirac Masses

There are two possibilities for neutrino mass:

Dirac Mass: Right-handed neutrino would be a new as-yet unseen type of particle. But it would interact so weakly that it would not have been produced in significant numbers during the big bang.

Majorana Mass: If *lepton number* is not conserved (which seems plausible), so the neutrino is absolutely neutral, then the right-handed neutrino could be the particle that we have called the anti-neutrino.

Neutrino Masses and Neutrinoless Double Beta Decay

☆ Key experiment to distinguish Majorana from Dirac mass: neutrinoless double beta decay. Standard double beta decay looks like

$$\boxed{(A, Z) \to (A, Z+2) + 2e^- + 2\bar{\nu}_e \ .}$$

If the ν has a Majorana mass, and therefore it is its own antiparticle, then the reaction could happen without the two final $\bar{\nu}_e$'s, which can essentially annihilate each other. (The annihilation could happen as part of the interaction, so the energy is given to the $(A, Z+2)$ and $2e^-$ particles, with no other particles emitted.)

8.286 Class 17
November 7, 2022

BLACK-BODY RADIATION AND THE EARLY HISTORY OF THE UNIVERSE, PART 4

Announcements: Quiz on Wednesday

- See web page for full information.
- Office hours:
  - Alan: today, 5:15-6:15 pm, Room 8-320.
  - Marianne: tomorrow, 4-5 pm in Room 6C-442 (also called the Cosman Room).
- Time and Place: The quiz will be during our regular class time on Wednesday, 11:05 am - 12:25 pm, in Room 50-340 (Walker Memorial).

The Real Story of Neutrino Masses

- We have not yet measured the mass of a neutrino, but we have seen neutrinos “oscillate” from one flavor to another:
  - Electron neutrinos from the Sun arrive at Earth as a mixture of all three flavors.
  - Neutrinos produced by cosmic rays in the upper atmosphere have been found to undergo oscillations on their way to ground level.
  - Neutrinos produced by reactors and accelerators have been seen to oscillate.
- Oscillations require a nonzero mass: essentially because a massless particle experiences an infinite time dilation, so time stops.
- The oscillations measure the differences of the squares of the masses.

Neutrino Masses and Quantum Superpositions

- Quantum theory allows for states that are superpositions of other states.
- Neutrinos are produced in states of definite flavor, called ν_e, ν_μ, and ν_τ. (I.e., $e^- + p \longrightarrow n + \nu_e$, while $\mu^- + p$ leads to ν_μ, and $\tau^- + p$ leads to ν_τ.) But these are not states of definite mass!
- The states of definite mass are called ν_1, ν_2, and ν_3.
- Each flavor state is a superposition of all three states of definite mass, and each state of definite mass is a superposition of all three flavor states.

Differences of Squares of Neutrino Masses

As of 2020, the Particle Data Group reports:

$$\Delta m_{21}^2\, c^4 = (7.53 \pm 0.18) \times 10^{-5}\ \mathrm{eV}^2\ ,$$

$$\Delta m_{32}^2\, c^4 = \left(2.546^{+0.034}_{-0.040}\right) \times 10^{-3}\,\mathrm{eV}^2\ ,$$

or

$$\Delta m_{32}^2\, c^4 = (2.453 \pm 0.034) \times 10^{-3}\,\mathrm{eV}^2\ ,$$

where the two options for Δm_{32}^2 depend on assumptions about the ordering of the masses. Note that $\sqrt{\Delta m_{21}^2\, c^4} = 8.68\times10^{-3}$ eV, and $\sqrt{\Delta m_{32}^2\, c^4} = 0.0505$ eV or 0.0495 eV. Recall that today $kT_\gamma = 2.35 \times 10^{-4}$ eV, which is much smaller.

Does Neutrino Mass Affect Our Calculation of $t_{\rm eq}$?

No!

We wrote

$$u_{\rm rad,0} = \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right]\frac{\pi^2}{30}\frac{(kT_\gamma)^4}{(\hbar c)^3}$$
$$= 7.01 \times 10^{-14}\ {\rm J/m^3}\ ,$$

but what we really used was

$$u_{\rm rad}(t) = \left[2 + \frac{21}{4}\left(\frac{4}{11}\right)^{4/3}\right]\frac{\pi^2}{30}\frac{(kT_\gamma)^4}{(\hbar c)^3}\left(\frac{a(t_0)}{a(t)}\right)^4\ ,$$

which is valid for t anywhere near the time $t_{\rm eq}$.

Cosmological Bound on the Sum of ν Masses

☆ From cosmology of large-scale structure, we know that*

$$\boxed{(m_1 + m_2 + m_3)c^2 \leq 0.17 \text{ eV}.}$$

☆ Why? Because neutrinos "free-stream" easily from one place to another. If they carried too much mass, they would even out the mass density and suppress large-scale structure.

*S. R. Choudhury and S. Hannestad, JCAP 2020, No. 7, 037 (2020), arXiv:1907.12598.

Neutrino Mass and Spin States

- ☆ The measurements of the mass differences imply that at least 2 of the 3 neutrino masses must be nonzero.
- ☆ If the mass of a neutrino is nonzero, then it **cannot** always be left-handed.
- ☆ To see this, consider a left-handed neutrino moving in the z direction, with spin in the $-z$ direction. With $m > 0$, it must move slower than c. So an observer can move along the z-axis faster than the neutrino. To such an observer, the momentum of the ν will be in the $-z$ direction, the spin will be in the $-z$ direction, and the ν will appear right-handed.
- ☆ How could this right-handed neutrino fit into our theory?

Majorana and Dirac Masses

There are two possibilities for neutrino mass:

Dirac Mass: Right-handed neutrino would be a new as-yet unseen type of particle. But it would interact so weakly that it would not have been produced in significant numbers during the big bang.

Majorana Mass: If *lepton number* is not conserved (which seems plausible), so the neutrino is absolutely neutral, then the right-handed neutrino could be the particle that we have called the anti-neutrino.

Neutrino Masses and Neutrinoless Double Beta Decay

☆ Key experiment to distinguish Majorana from Dirac mass: neutrinoless double beta decay. Standard double beta decay looks like

$$\boxed{(A, Z) \to (A, Z+2) + 2e^- + 2\bar{\nu}_e \ .}$$

If the ν has a Majorana mass, and therefore it is its own antiparticle, then the reaction could happen without the two final $\bar{\nu}_e$'s, which can essentially annihilate each other. (The annihilation could happen as part of the interaction, so the energy is given to the $(A, Z+2)$ and $2e^-$ particles, with no other particles emitted.)

Thermal History of the Universe

Assuming that the early universe can be described as radiation-dominated and flat (excellent approximations), then

$$H^2 = \frac{8\pi}{3} G\rho \ , \quad a(t) \propto t^{1/2} \ , \quad H = \frac{\dot{a}}{a} = \frac{1}{2t} \ ,$$

which implies

$$\boxed{\rho = \frac{3}{32\pi G t^2} \ .}$$

We also know

$$u = g \frac{\pi^2}{30} \frac{(kT)^4}{(\hbar c)^3} \ , \quad \text{and} \quad \rho = u/c^2 \ ,$$

so

$$\boxed{kT = \left(\frac{45\hbar^3 c^5}{16\pi^3 g G} \right)^{1/4} \frac{1}{\sqrt{t}} \ .}$$

$$kT = \left(\frac{45\hbar^3 c^5}{16\pi^3 g G}\right)^{1/4} \frac{1}{\sqrt{t}} \,.$$

Assuming $0.511\text{ MeV} \ll kT \ll 106\text{ MeV}$ (i.e., assuming kT is between mc^2 for the electron and muon),

$$g_{\rm tot} = \underbrace{2}_{\text{photons}} + \underbrace{\frac{21}{4}}_{\text{neutrinos}} + \underbrace{\frac{7}{2}}_{e^+e^-} = 10\frac{3}{4} \,.$$

For $t = 1$ second, this gives $kT = 0.860$ MeV.

Assuming $0.511 \text{ MeV} \ll kT \ll 106 \text{ MeV}$ (i.e., assuming kT is between mc^2 for the electron and muon), we find that at $t = 1$ second, $kT = 0.860$ MeV.

Since $T \propto 1/\sqrt{t}$, we can write

$$\boxed{kT = \frac{0.860 \text{ MeV}}{\sqrt{t \text{ (in sec)}}} \quad ,}$$

or equivalently

$$\boxed{T = \frac{9.98 \times 10^9 \text{ K}}{\sqrt{t \text{ (in sec)}}} \quad .}$$

Relation Between a and T

☆ Conservation of entropy implies that $s \propto 1/a^3(t)$, but we also know that $s \propto gT^3$. It follows that

$$\boxed{g^{1/3}T \propto \frac{1}{a(t)} .}$$

Recombination

- "Baryonic" matter is matter made of protons, neutrons, and electrons. I.e., it is ordinary matter, as opposed to dark matter or dark energy.
- About 80% of baryonic matter is hydrogen. Most of the rest is helium. Elements heavier then helium make up a very small fraction. So we mostly have hydrogen.
- At high T, hydrogen atoms ionize, become free protons and electrons. The ionization temperature depends on density, but for the density of the early universe, it is about 4,000 K. (Ryden calculates it on p. 154 as 3760 K.)
- When T falls below 4,000 K, the protons and electrons combine to form neutral H. This is called "recombination," but "combination" would be more accurate.

Decoupling

- Photons interact strongly with free electrons.

- The reason can be understood classically: when an electromagnetic wave hits a free electron, the electron experiences the $\vec{F} = e\vec{E}$ force of the electric field. Since its mass is very small, it oscillates rapidly, and sends electromatic radiation in all directions, using energy that it removes from the incoming wave. Thus, the incoming wave is scattered.

- The result is that the universe was opaque to photons in the ionized phase (plasma phase), but became transparent when the ionized gas became neutral atoms.

- The transition to a transparent universe is called "decoupling" (i.e., the photons "decouple" from the matter of the universe).

Time of Decoupling t_d

- ☆ The result is that the universe was opaque to photons in the ionized phase (plasma phase), but became transparent when the ionized gas became neutral atoms.
- ☆ The transition to a transparent universe is called "decoupling" (i.e., the photons "decouple" from the matter of the universe).

- ☆ At $T_{\rm rec} = 4,000$ K, about half of the hydrogen is ionized.
- ☆ Note that $kT_{\rm rec} \approx 0.34$ eV, while the ionization energy of H is 13.6 eV. The reason for the big difference is that $n_{\rm baryon}/n_\gamma \approx 10^{-9}$, so it is rare for electrons and protons to find each other.
- ☆ Since even a very small density of free electrons is enough to make the universe opaque, photon decoupling does not occur until T falls to $T_{\rm dec} \approx 3,000$ K.

- Since even a very small density of free electrons is enough to make the universe opaque, photon decoupling does not occur until T falls to $T_{\rm dec} \approx 3,000$ K.

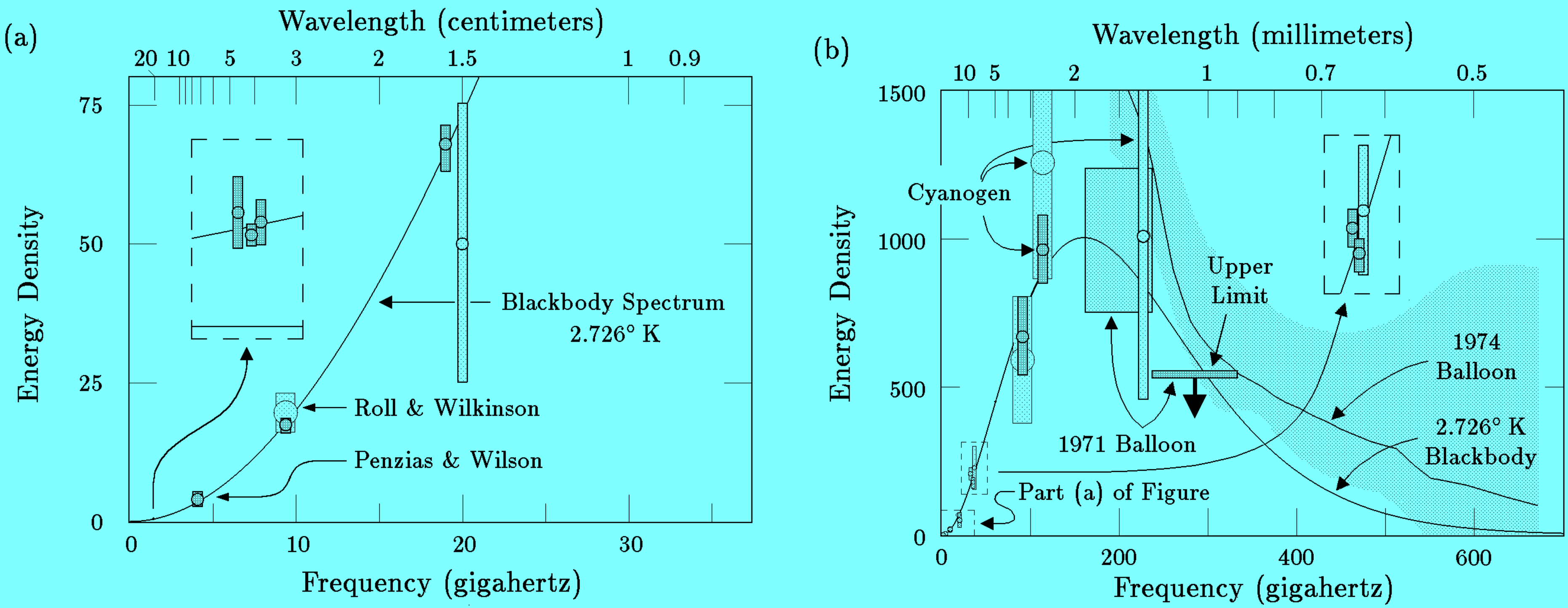
- Approximating the universe as flat and matter-dominated from $T_{\rm dec}$ to today, we can estimate the time of decoupling by

$$\boxed{\begin{aligned} t_d &= \left(\frac{T_0}{T_d}\right)^{3/2} t_0 \\ &\approx \left(\frac{2.7\,\mathrm{K}}{3000\,\mathrm{K}}\right)^{3/2} \times \left(13.7\times 10^9 \text{ yr}\right) \approx 370,000 \text{ yr} . \end{aligned}}$$

Spectrum of the Cosmic Microwave Background

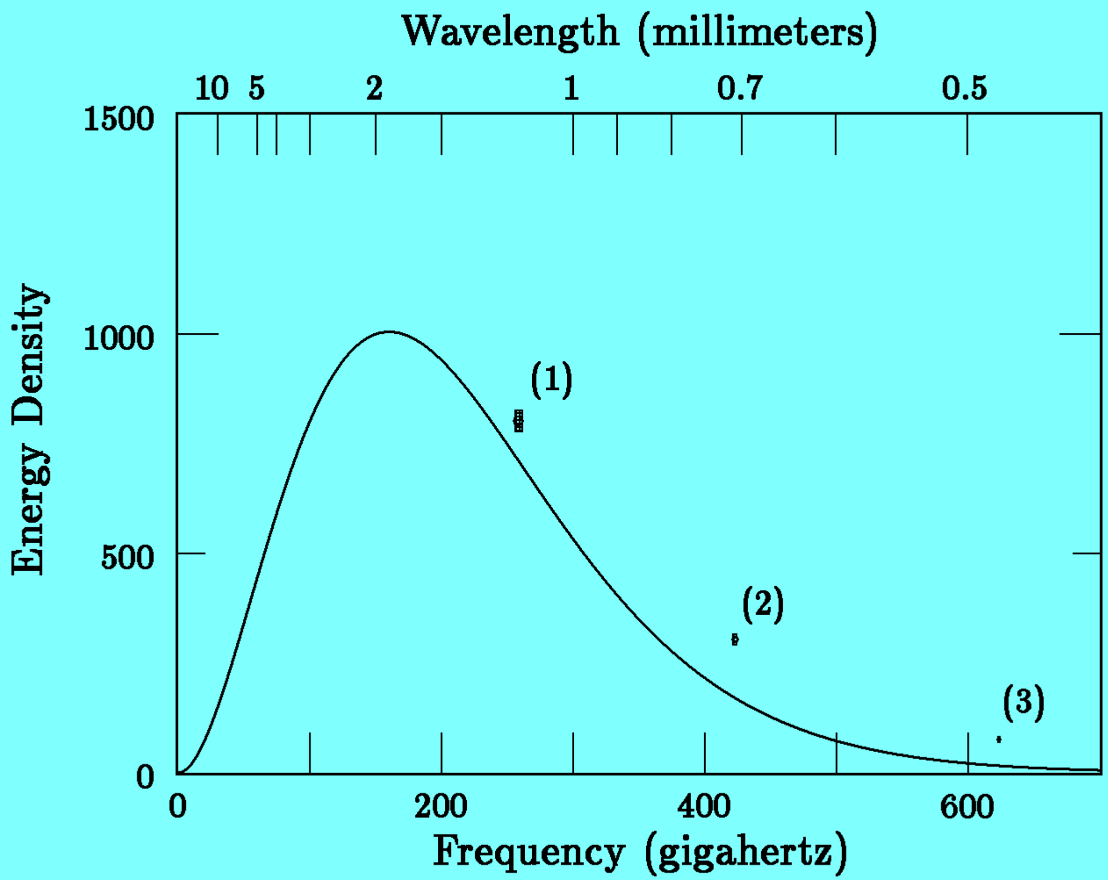
$$\rho_\nu(\nu)\,\mathrm{d}\nu = \frac{16\pi^2\hbar\nu^3}{c^3}\frac{1}{e^{2\pi\hbar\nu/kT}-1}\,\mathrm{d}\nu\ ,$$

where $\rho_\nu(\nu)\,\mathrm{d}\nu$ is the energy density of photons in the frequency range from ν to $\nu + \mathrm{d}\nu$.



CMB Data in 1975

The situation got worse before it got better:



Data from Berkeley-Nagoya Rocket Flight, 1987

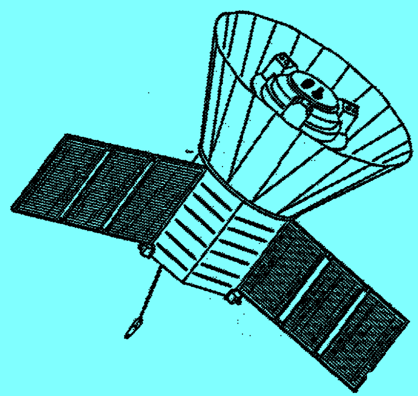
Points 2 and 3 differ from the blackbody curve by 12 and 16 standard deviations, respectively!

Preprint No. 90–01

COBE PREPRINT

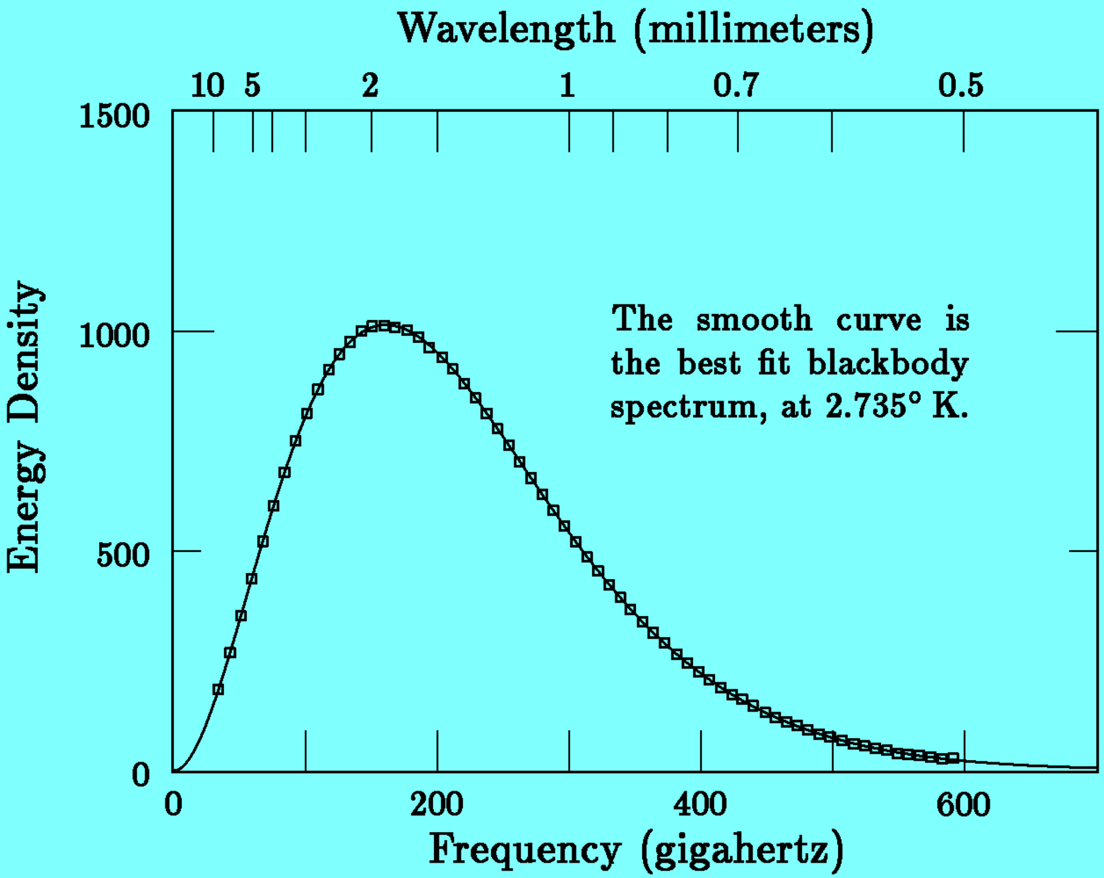
A PRELIMINARY MEASUREMENT OF THE COSMIC MICROWAVE BACKGROUND SPECTRUM BY THE COSMIC BACKGROUND EXPLORER (COBE) SATELLITE

J.C. Mather, E. S. Cheng, R. E. Eplee, R. B. Isaacman, S. S. Meyer, R. A. Shafer, R. Weiss, E. L. Wright, C. L. Bennett, N. W. Boggess, E. Dwek, S. Gulkis, M. G. Hauser, M. Janssen, T. Kelsall, P. M. Lubin, S. H. Moseley, Jr., T. L. Murdock, R. F. Silverberg, G. F. Smoot, and D. T. Wilkinson.



COSMIC BACKGROUND EXPLORER

The Cosmic Background Explorer (COBE) satellite was launched in the fall of 1989. In January 1990, at meeting of the American Astronomical Society in Washington, D.C., the first data on the spectrum of the cosmic microwave background was announced. Shown is the cover page of the original preprint.



This is the original COBE measurement of the CMB spectrum, Jan 1990. The Energy density is in units of electron volts per cubic meter per gigahertz. The error bars are shown as 1% of the peak intensity. This graph was based on 9 minutes of data. Later data analysis reduced the error bars by a factor of 100, with still a perfect fit to the blackbody spectrum.

Historical Interlude:

Albert Einstein
and
Alexander Friedmann

Albert Einstein and the Friedmann Equations

Albert Einstein

Alexander A. Friedmann

Publication of the Friedmann Equations

On the Curvature of Space

A. Friedmann
Petersburg
Received June 29, 1922
Zeitschrift für Physik

Einstein's Reaction

REMARK ON THE WORK OF A. FRIEDMANN (FRIEDMANN 1922)
"ON THE CURVATURE OF SPACE"

A. Einstein, Berlin

Received September 18, 1922
Zeitschrift für Physik

The work cited contains a result concerning a non-stationary world which seems suspect to me. Indeed, those solutions do not appear compatible with the field equations (A). From the field equations it follows necessarily that the divergence of the matter tensor T_{ik} vanishes. This along with the anzatzes (C) and (D) leads to the condition

$$\partial\rho/\partial x_4 = 0$$

which together with (8) implies that the world-radius R is constant in time. The significance of the work therefore is to demonstrate this constancy.

REFERENCES: Friedmann, A. 1922, *Zs. f. Phys.*, 10, 377.

Translation: *Cosmological Constants*, edited by Jeremy Bernstein and Gerald Feinberg

Sequence of Events

June 29, 1922: Friedmann's paper received at *Zeitschrift für Physik.*

September 18, 1922: Einstein's refutation received at *Zeitschrift für Physik.*

December 6, 1922: Friedmann learns about Einstein's objection from his friend, Yuri A. Krutkov, who is visiting in Berlin. Friedmann writes a detailed letter to Einstein. Einstein is traveling and does not read it.

May, 1923: Einstein meets Krutkov in Leiden, both attending the farewell lecture by Lorentz, who was retiring.

Krutkov's letters to his sister: "On Monday, May 7, 1923, I was reading, together with Einstein, Friedmann's article in the *Zeitschrift für Physik.*" May 18: "I defeated Einstein in the argument about Friedmann. Petrograd's honor is saved!"*

May 31, 1923: Einstein's retraction of his refutation is received at *Zeitschrift für Physik.*

* Quoted in *Alexander A. Friedmann: the Man who Made the Universe Expand*, by E.A. Tropp, V. Ya. Frenkel, & A.D. Chernin.

Einstein's Retraction

A NOTE ON THE WORK OF A. FRIEDMANN
"ON THE CURVATURE OF SPACE"

A. Einstein, Berlin

Received May 31, 1923
Zeitschrift für Physik

I have in an earlier note (Einstein 1922) criticized the cited work (Friedmann 1922). My objection rested however — as Mr. Krutkoff in person and a letter from Mr. Friedmann convinced me — on a calculational error. I am convinced that Mr. Friedmann's results are both correct and clarifying. They show that in addition to the static solutions to the field equations there are time varying solutions with a spatially symmetric structure.

REFERENCES:

Einstein, A. 1922, *Zs. f. Phys.*, 11, 326.
Friedmann, A. 1922, *Ebenda*, 10, 377.

Translation: *Cosmological Constants*, edited by Jeremy Bernstein and Gerald Feinberg

Einstein and Krutkov

Albert Einstein
Barcelona, 1923

Yuri A. Krutkov.

Einstein's Draft

Notiz zu der Arbeit von A. Friedmann
„Über die Krümmung des Raumes"

Ich habe in einer früheren Notiz[x] an der genannten Arbeit[xx] Kritik geübt. Mein Einwand beruhte aber – wie ich ~~mich auf Anregung von Herrn Krutkoff~~ (an Hand eines Briefes von Herrn Friedmann) überzeugt habe – auf einem Rechenfehler. Ich halte Herrn ~~Krut~~ Friedmanns Resultate für richtig und ~~interessant~~ aufklärend. Es zeigt sich, dass die Feldgleichungen ~~dyn~~ neben den statischen dynamische (zentrisch-symmetrische) (d. h. mit der Zeitkoordinate veränderliche) für die Raumstruktur Lösungen zulassen, ~~denen eine physikalische Bedeutung kaum zuzuschreiben sein dürfte~~.

A. Einstein.

x Zeitschr. für Physik 1922 11.B. S. 326

xx Zeitschr. für Physik 1922 10.B S. 322.

Einstein's draft of 1923 in which he withdrew his earlier objection to Friedmann's dynamic solutions to the field equations. The last bit of the last sentence was: "a physical significance can hardly be ascribed to them". He crossed this out before sending the note to print.

"a physical significance can hardly be ascribed to them."

* From *The Invented Universe*, by Pierre Kerszberg

A Brief History of the Cosmological Constant

- ☆ In 1917, Einstein applied his new GR to the universe, and discovered that a static universe would collapse.

- ☆ Convinced that the universe was static, Einstein introduced the *cosmological constant* Λ into his field equations — the equations that describe how matter affects the metric — to create a gravitational repulsion to oppose the collapse.

- ☆ From a modern point of view, Λ represents a *vacuum energy density* u_{vac}, with

$$\boxed{u_{\text{vac}} = \rho_{\text{vac}} c^2 = \frac{\Lambda c^4}{8\pi G}\ ,}$$

  because u_{vac} appears in the field equations exactly as a vacuum energy density would. To Einstein, however, it was simply a new term in the field equations. Before quantum theory, the vacuum was viewed as completely empty, so it was inconceivable that it could have a nonzero energy density.

- Once the expansion of the universe was discovered by Hubble in 1929, Einstein abandoned Λ as being no longer needed or wanted.
- In 1998, however, two (large) groups of astronomers, both using measurements of Type Ia supernovae at redshifts $z \lesssim 1$, discovered evidence that the expansion of the universe is currently accelerating. At the time, it was shocking! *Science* magazine proclaimed it (correctly!) as the "Breakthough of the Year".
- In 2011 the Nobel Prize in Physics was awarded to Saul Permutter, Brian Schmidt, and Adam Riess for this discovery. In 2015 the Breakthrough Prize in Fundamental Physics was awarded to these three, and also the two entire teams.

8.286 Class 20
November 16, 2022

THE COSMOLOGICAL CONSTANT

Announcements

☆ Office hours, Friday at 1:00 pm, Room 4-144. I'll be taking over Marianne's office hour this Friday.

A Brief History of the Cosmological Constant

- In 1917, Einstein applied his new GR to the universe, and discovered that a static universe would collapse.
- Convinced that the universe was static, Einstein introduced the *cosmological constant* Λ into his field equations — the equations that describe how matter affects the metric — to create a gravitational repulsion to oppose the collapse.
- From a modern point of view, Λ represents a *vacuum energy density* $u_{\rm vac}$, with

$$\boxed{u_{\rm vac} = \rho_{\rm vac} c^2 = \frac{\Lambda c^4}{8\pi G}\ ,}$$

because $u_{\rm vac}$ appears in the field equations exactly as a vacuum energy density would. To Einstein, however, it was simply a new term in the field equations. Before quantum theory, the vacuum was viewed as completely empty, so it was inconceivable that it could have a nonzero energy density.

- Once the expansion of the universe was discovered by Hubble in 1929, Einstein abandoned Λ as being no longer needed or wanted.
- In 1998, however, two (large) groups of astronomers, both using measurements of Type Ia supernovae at redshifts $z \lesssim 1$, discovered evidence that the expansion of the universe is currently accelerating. At the time, it was shocking! *Science* magazine proclaimed it (correctly!) as the "Breakthough of the Year".
- In 2011 the Nobel Prize in Physics was awarded to Saul Permutter, Brian Schmidt, and Adam Riess for this discovery. In 2015 the Breakthrough Prize in Fundamental Physics was awarded to these three, and also the two entire teams.

Gravitational Effect of Pressure

$$\frac{d^2a}{dt^2} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a \ .$$

Gravitational Effect of Pressure

$$\frac{d^2a}{dt^2} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a \ .$$

Vacuum Energy and the Cosmological Constant:

$$u_{\text{vac}} = \rho_{\text{vac}}c^2 = \frac{\Lambda c^4}{8\pi G} \ .$$

Gravitational Effect of Pressure

$$\frac{d^2a}{dt^2} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right)a \ .$$

Vacuum Energy and the Cosmological Constant:

$$u_{\rm vac} = \rho_{\rm vac}c^2 = \frac{\Lambda c^4}{8\pi G} \ .$$

Recall that

$$\dot\rho = -3\frac{\dot a}{a}\left(\rho + \frac{p}{c^2}\right) \ ,$$

where the overdot indicates a time derivative. So

$$\dot\rho_{\rm vac} = 0 \quad \Longrightarrow \quad p_{\rm vac} = -\rho_{\rm vac}c^2 = -\frac{\Lambda c^4}{8\pi G} \ .$$

Defining $\boxed{\rho = \rho_n + \rho_{\text{vac}}}$ and $\boxed{p = p_n + p_{\text{vac}},}$ the Friedmann equations become:

$$\boxed{\ddot{a} = -\frac{4\pi}{3} G \left(\rho_n + \frac{3p_n}{c^2} - 2\rho_{\text{vac}} \right) a \ .}$$

$$\boxed{\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi}{3} G (\rho_n + \rho_{\text{vac}}) - \frac{kc^2}{a^2} \ ,}$$

where an overdot ($\dot{\ }$) is a derivative with respect to t (cosmic time).

Defining $\boxed{\rho = \rho_n + \rho_{\text{vac}}}$ and $\boxed{p = p_n + p_{\text{vac}},}$ the Friedmann equations become:

$$\boxed{\ddot{a} = -\frac{4\pi}{3} G \left(\rho_n + \frac{3p_n}{c^2} - 2\rho_{\text{vac}} \right) a \ .}$$

$$\boxed{\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi}{3} G (\rho_n + \rho_{\text{vac}}) - \frac{kc^2}{a^2} \ ,}$$

where an overdot ($\dot{}$) is a derivative with respect to t (cosmic time). At late times, $\rho_n \propto 1/a^3$ or $1/a^4$, $\rho_{\text{vac}} = \text{constant}$, so if $a(t)$ keeps growing, then ρ_{vac} dominates. Then

$$\boxed{\begin{gathered} a(t) \propto e^{H_{\text{vac}} t} \ , \\ H \to H_{\text{vac}} = \sqrt{\frac{8\pi}{3} G \rho_{\text{vac}}} \ . \end{gathered}}$$

Age of the Universe with Λ

The first order Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3} G \left(\underbrace{\rho_m}_{\propto \frac{1}{a^3(t)}} + \underbrace{\rho_{\rm rad}}_{\propto \frac{1}{a^4(t)}} + \rho_{\rm vac} \right) - \frac{kc^2}{a^2} .$$

can be rewritten as

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left(\frac{\Omega_{m,0}}{x^3} + \frac{\Omega_{\rm rad,0}}{x^4} + \Omega_{\rm vac} \right) - \frac{kc^2}{a^2} ,$$

$$\text{where} \quad x \equiv a(t)/a(t_0) .$$

and where we used

$$\Omega_{X,0} = \frac{\rho_{X,0}}{\rho_{c,0}} = \frac{8\pi G \rho_{X,0}}{3H_0^2} .$$

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left(\frac{\Omega_{m,0}}{x^3} + \frac{\Omega_{\mathrm{rad},0}}{x^4} + \Omega_{\mathrm{vac}}\right) - \frac{kc^2}{a^2}\ ,$$

$$\text{where}\quad x \equiv a(t)/a(t_0)\ .$$

Define

$$\Omega_{k,0} \equiv -\frac{kc^2}{a^2(t_0)H_0^2}\ .$$

So

$$\left(\frac{\dot{a}}{a}\right)^2 = \left(\frac{\dot{x}}{x}\right)^2 = \frac{H_0^2}{x^4}\left(\Omega_{m,0}x + \Omega_{\mathrm{rad},0} + \Omega_{\mathrm{vac},0}x^4 + \Omega_{k,0}x^2\right)\ .$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \left(\frac{\dot{x}}{x}\right)^2 = \frac{H_0^2}{x^4}\left(\Omega_{m,0}x + \Omega_{\mathrm{rad},0} + \Omega_{\mathrm{vac},0}x^4 + \Omega_{k,0}x^2\right) \ .$$

At present time, $\dot{a}/a = H_0$ and $x = 1$, so the sum of the Ω's must equal 1. Thus, $\Omega_{k,0}$ can be evaluated from

$$\Omega_{k,0} = 1 - \Omega_{m,0} - \Omega_{\mathrm{rad},0} - \Omega_{\mathrm{vac},0} \ .$$

Observationally, $\Omega_{k,0}$ is consistent with zero, but we can still allow for it in our final formula for the age:

$$t_0 = \frac{1}{H_0}\int_0^1 \frac{x\,dx}{\sqrt{\Omega_{m,0}x + \Omega_{\mathrm{rad},0} + \Omega_{\mathrm{vac},0}x^4 + \Omega_{k,0}x^2}} \ .$$

Numerical Integration with Mathematica

Reference: N. Aghanim et al. (Planck Collaboration), “Planck 2018 results, VI: Cosmological parameters,” Table 2, Column 6, arXiv:1807.06209.

IN: t0[H0_,Ωm0_,Ωrad0_,Ωvac0_,Ωk0_] := (1/H0) *
NIntegrate[x/Sqrt[Ωm0 x + Ωrad0 + Ωvac0 x^4 + Ωk0 x^2], {x,0,1}]

IN: PlanckH0 := Quantity[67.66,"km/sec/Mpc"]

IN: PlanckΩm0 := 0.3111

IN: PlanckΩvac0 := 0.6889

IN: Ωrad0 := $4.15 \times 10^{-5}\, h_0^{-2} = 9.07 \times 10^{-5}$

IN: UnitConvert[t0[PlanckH0, PlanckΩm0 − Ωrad0/2, Ωrad0, PlanckΩvac0 − Ωrad0/2 ,0], "Years"]

OUT: 1.3796×10^{10} years

The Planck paper gives 13.787 ± 0.020 Gyr. The difference is about 9 million years, 0.06%, or 0.45 σ.

Look-Back Time

Question: If we observe a distant galaxy at redshift z, how long has it been since the light left the galaxy? The answer is called the *look-back time*.

To answer, recall that we wrote t_0 as an integral over $x = a(t)/a(t_0)$. We can change variables to

$$\boxed{1 + z = \frac{a(t_0)}{a(t)} = \frac{1}{x} \ ,}$$

which gives

$$t_0 = \frac{1}{H_0} \int_0^\infty \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\text{rad},0}(1+z)^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z)^2}} \ .$$

$$t_0 = \frac{1}{H_0}\int_0^\infty \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\mathrm{rad},0}(1+z)^4 + \Omega_{\mathrm{vac},0} + \Omega_{k,0}(1+z)^2}} \ .$$

The integral over any interval of z gives the corresponding time interval, so the look-back time is just the integral from 0 to z:

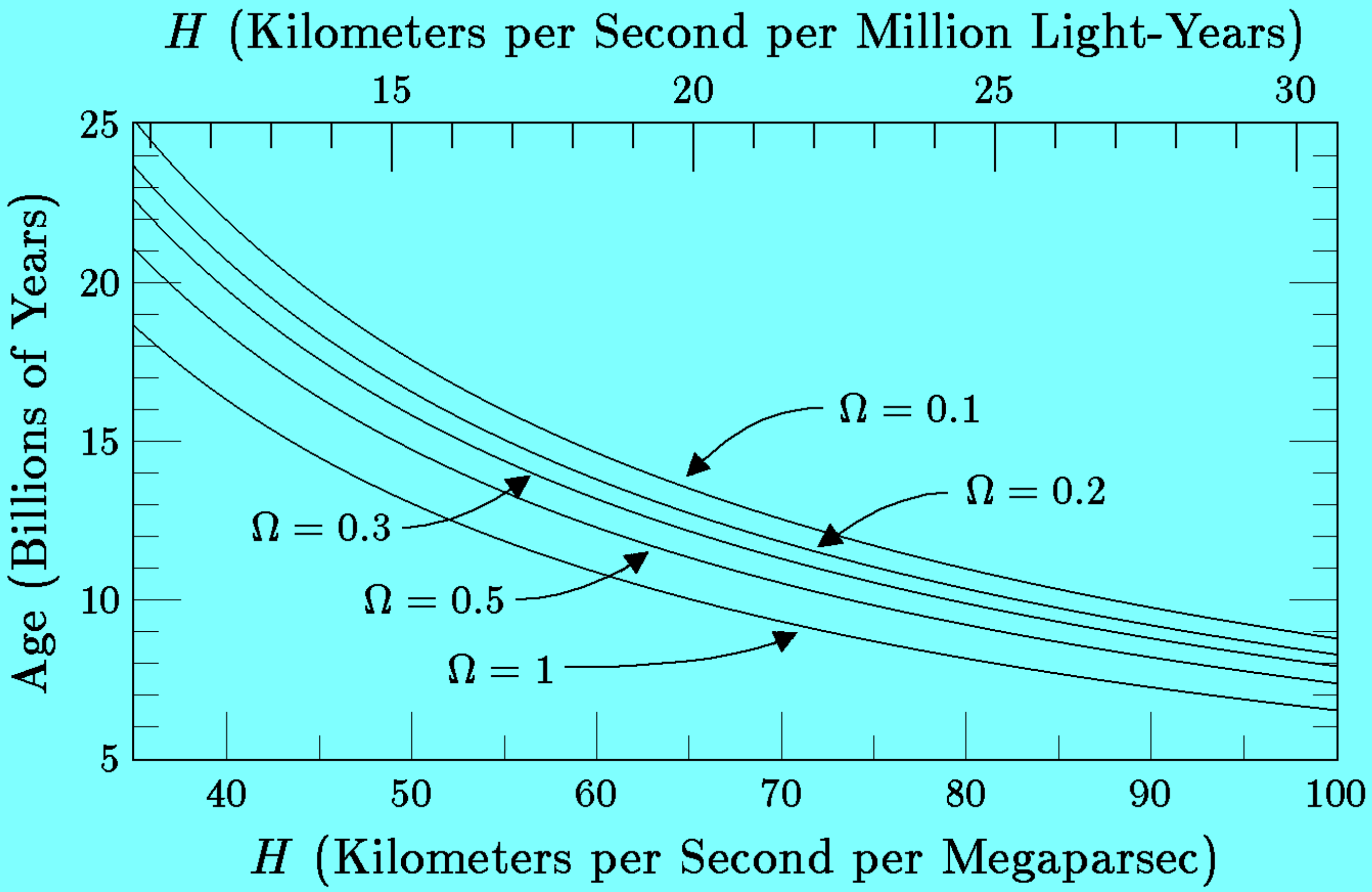
$$t_{\text{look-back}}(z) = \frac{1}{H_0}\int_0^z \frac{dz'}{(1+z')\sqrt{\Omega_{m,0}(1+z')^3 + \Omega_{\mathrm{rad},0}(1+z')^4 + \Omega_{\mathrm{vac},0} + \Omega_{k,0}(1+z')^2}} \ .$$

Age of a Flat Universe with Λ and Matter Only

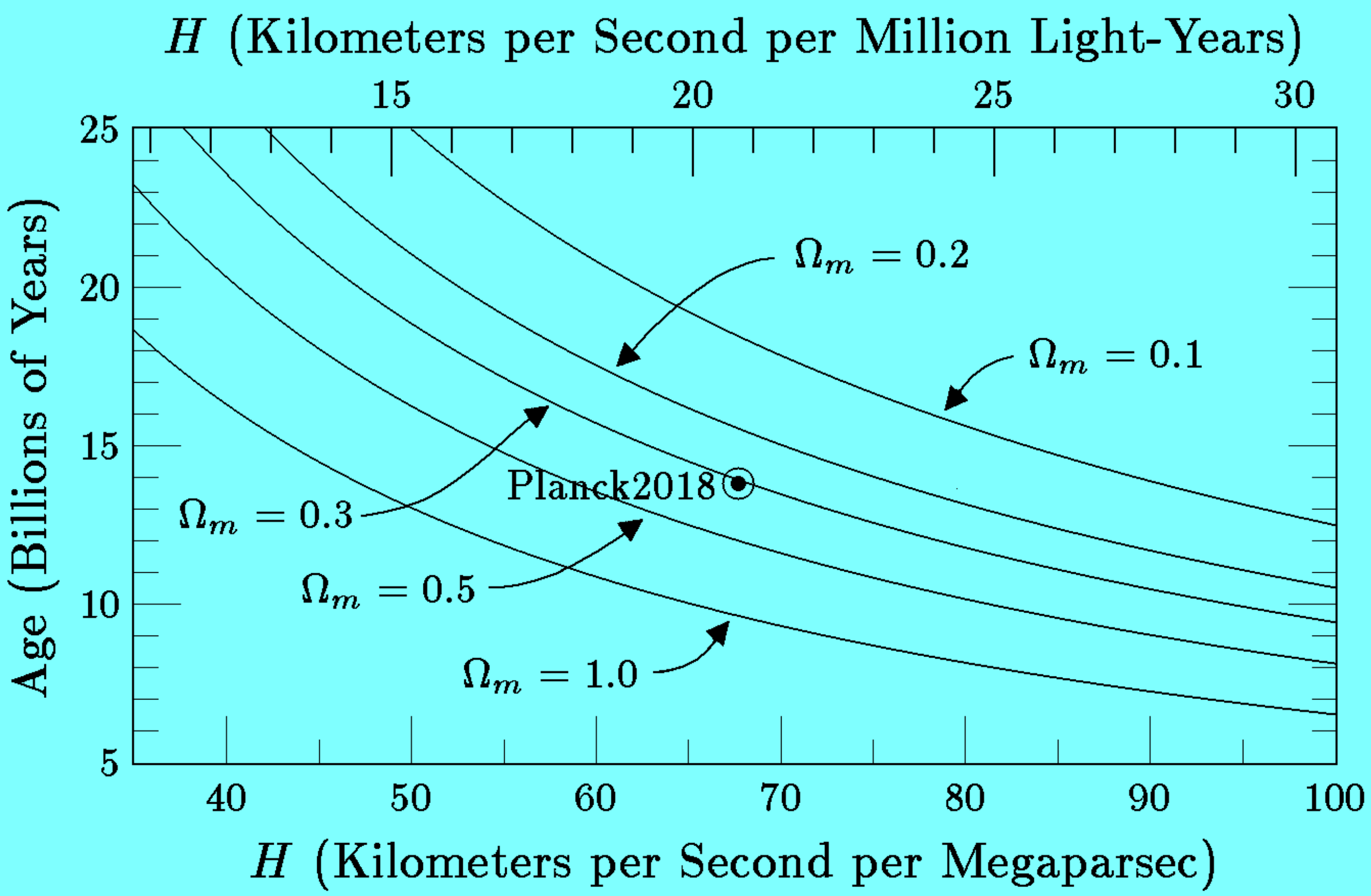
If $\Omega_{\text{rad}} = \Omega_k = 0$, then it is possible to carry out the integral for the age analytically:

$$t_0 = \begin{cases} \dfrac{2}{3H_0} \dfrac{\tan^{-1}\sqrt{\Omega_{m,0}-1}}{\sqrt{\Omega_{m,0}-1}} & \text{if } \Omega_{m,0} > 1,\ \Omega_{\text{vac}} < 0 \\ \dfrac{2}{3H_0} & \text{if } \Omega_{m,0} = 1,\ \Omega_{\text{vac}} = 0 \\ \dfrac{2}{3H_0} \dfrac{\tanh^{-1}\sqrt{1-\Omega_{m,0}}}{\sqrt{1-\Omega_{m,0}}} & \text{if } \Omega_{m,0} < 1,\ \Omega_{\text{vac}} > 0\ . \end{cases}$$

The Age Problem with Only Nonrelativistic Matter



Age of a Flat Universe with Λ and Matter Only



Ryden Benchmark and Planck 2018 Best Fit

Parameters	Ryden Benchmark	Planck 2018 Best Fit
$\boldsymbol{H_0}$	68	$67.7 \pm 0.4\ \mathrm{km{\cdot}s^{-1}{\cdot}Mpc^{-1}}$
Baryonic matter $\boldsymbol{\Omega_b}$	0.048	$0.0490 \pm 0.0007^*$
Dark matter $\boldsymbol{\Omega_{\mathrm{dm}}}$	0.262	$0.261 \pm 0.004^*$
Total matter $\boldsymbol{\Omega_m}$	0.31	0.311 ± 0.006
Vacuum energy $\boldsymbol{\Omega_{\mathrm{vac}}}$	0.69	0.689 ± 0.006

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54 \text{ km sec}^{-1} \text{ Mpc}^{-1}$

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54$ km sec^{-1} Mpc^{-1}

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04$ km sec^{-1} Mpc^{-1}.

Controversy in Parameters: “Hubble Tension”

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54$ km sec^{-1} Mpc^{-1}

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04$ km sec^{-1} Mpc^{-1}.

- The difference is about 4.85 σ. If the discrepancy is random and the normal probability distribution applies, the probability of such a large deviation is about 1 in 800,000.

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54$ km sec^{-1} Mpc^{-1}

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04$ km sec^{-1} Mpc^{-1}.

- The difference is about 4.85 σ. If the discrepancy is random and the normal probability distribution applies, the probability of such a large deviation is about 1 in 800,000.

- From the "tip of the red giant branch",

  Wendy Freedman's group: $H_0 = 69.8 \pm 0.6(\text{stat}) \pm 1.6(\text{sys})$ km sec^{-1} Mpc^{-1}

References: Planck Collaboration, Astron. & Astrophys. 641, A6 (2020) [arXiv:1807.06209v4], Table 2, Col. 6

A. Riess et al., Astrophys. J. 934, L7 (2022) [arXiv:2112.04510].

W. Freedman, Astrophys. J. 919, 16 (2021) [arXiv:2106.15656].

The Hubble Diagram: Radiation Flux vs. Redshift

If we live in a universe like we have described, what do we expect to find if we measure the energy flux from a "standard candle" as a function of its redshift?

Consider closed universe:

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} .$$

We will be interested in tracing radial trajectories, so we can simplify the radial metric by a change of variables

$$\boxed{\sin\psi \equiv \sqrt{k}\,r\ .}$$

Then

$$d\psi = \frac{\sqrt{k}\,dr}{\cos\psi} = \frac{\sqrt{k}\,dr}{\sqrt{1-kr^2}} ,$$

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\}\ .$$

$$d\psi = \frac{\sqrt{k}\,dr}{\cos\psi} = \frac{\sqrt{k}\,dr}{\sqrt{1-kr^2}}\ ,$$

---

and the metric simplifies to

$$\boxed{ds^2 = -c^2\,dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\}\ ,}$$

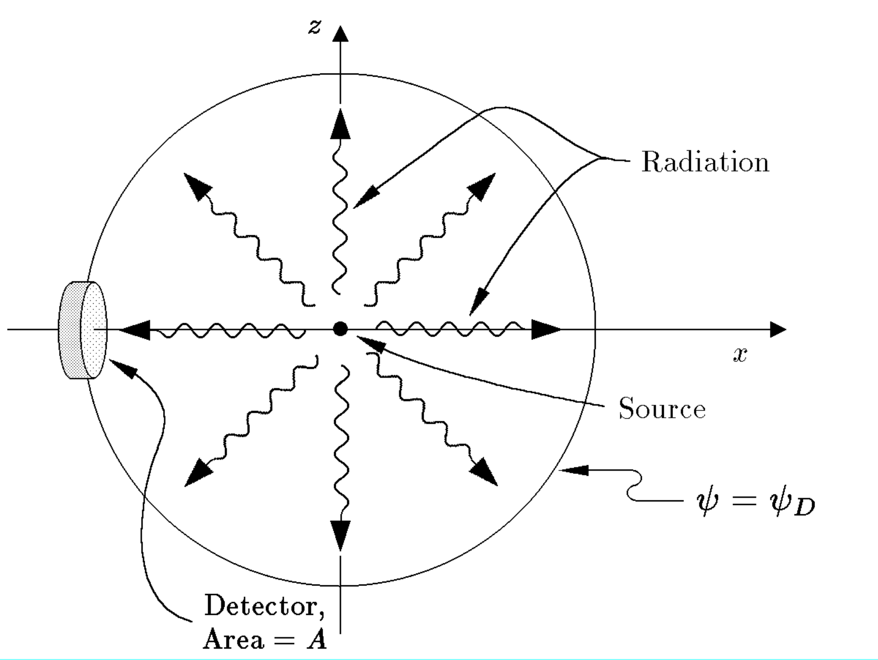
where

$$\tilde{a}(t) \equiv \frac{a(t)}{\sqrt{k}}\ .$$

Note: ψ is in fact the same angle ψ that we used in our construction of the closed-universe metric: it is the angle from the w-axis.

Geometry of Flux Calculation

$$ds^2 = -c^2\,dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\}$$



The fraction of the photons hitting the sphere that hit the detector is just the ratio of the areas:

$$\text{fraction} = \frac{\text{area of detector}}{\text{area of sphere}} = \frac{A}{4\pi\tilde{a}^2(t_0)\sin^2\psi_D}\ .$$

- The power hitting the sphere is less than the power P emitted by the source by two factors of $(1+z_S)$, where z_S is the redshift of the source: one factor due to redshift of each photon, and one factor due to the redshift of the rate of arrival of photons.

$$\boxed{P_{\text{received}} = \frac{P}{(1+z_S)^2}\frac{A}{4\pi\tilde{a}^2(t_0)\sin^2\psi_D} \ .}$$

Flux $J = P_{\text{received}}/A$.

8.286 Class 21
November 21, 2022

THE COSMOLOGICAL CONSTANT, PART 2

Age of the Universe

☆ Define $\Omega_{k,0}$ by

$$\Omega_{k,0} \equiv -\frac{kc^2}{a^2(t_0)H_0^2}$$
$$= 1 - \Omega_{m,0} - \Omega_{\text{rad},0} - \Omega_{\text{vac},0} \ .$$

where

$$\Omega_{m,0} \equiv \rho_{m,0}/\rho_{c,0} \ , \ \ \rho_{m,0} = \text{density of matter, today}$$
$$\Omega_{\text{rad},0} \equiv \rho_{\text{rad},0}/\rho_{c,0} \ , \ \ \rho_{\text{rad},0} = \text{density of radiation, today}$$
$$\Omega_{\text{vac},0} \equiv \rho_{\text{vac},0}/\rho_{c,0} \ , \ \ \rho_{\text{vac},0} = \text{density of vacuum energy, today}$$
$$\rho_{c,0} \equiv \text{critical density, today}$$
$$H_0 \equiv \text{Hubble expansion rate, today}$$

$$\Omega_{k,0} \equiv -\frac{kc^2}{a^2(t_0)H_0^2}$$
$$= 1 - \Omega_{m,0} - \Omega_{\text{rad},0} - \Omega_{\text{vac},0} \; .$$

☆ Then the age of the universe can be calculated as

$$t_0 = \frac{1}{H_0}\int_0^1 \frac{x\,dx}{\sqrt{\Omega_{m,0}x + \Omega_{\text{rad},0} + \Omega_{\text{vac},0}x^4 + \Omega_{k,0}x^2}}$$
$$= \frac{1}{H_0}\int_0^\infty \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\text{rad},0}(1+z)^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z)^2}} \; .$$

where $x \equiv a(t)/a(t_0)$, $1 + z = 1/x$.

Look-Back Time

$$t_0 = \frac{1}{H_0} \int_0^\infty \frac{dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\mathrm{rad},0}(1+z)^4 + \Omega_{\mathrm{vac},0} + \Omega_{k,0}(1+z)^2}} \; .$$

Question: If we observe a distant galaxy at redshift z, how long has it been since the light left the galaxy? The answer is called the *look-back time*.

$$t_{\text{look-back}}(z) = \frac{1}{H_0} \int_0^z \frac{dz'}{(1+z')\sqrt{\Omega_{m,0}(1+z')^3 + \Omega_{\mathrm{rad},0}(1+z')^4 + \Omega_{\mathrm{vac},0} + \Omega_{k,0}(1+z')^2}} \; .$$

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54 \text{ km sec}^{-1} \text{ Mpc}^{-1}$

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54$ km sec^{-1} Mpc^{-1}

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04$ km sec^{-1} Mpc^{-1}.

Controversy in Parameters: “Hubble Tension”

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54 \text{ km sec}^{-1} \text{ Mpc}^{-1}$

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04 \text{ km sec}^{-1} \text{ Mpc}^{-1}$.

- The difference is about 4.85 σ. If the discrepancy is random and the normal probability distribution applies, the probability of such a large deviation is about 1 in 800,000.

Controversy in Parameters: "Hubble Tension"

- From the CMB, the best number is from

  Planck 2018: $H_0 = 67.36 \pm 0.54$ km sec^{-1} Mpc^{-1}

- From standard candles and Cepheid variables,

  SH0ES (Supernovae, H0, for the Equation of State of dark energy, group led by Adam Riess): $H_0 = 73.04 \pm 1.04$ km sec^{-1} Mpc^{-1}.

- The difference is about 4.85 σ. If the discrepancy is random and the normal probability distribution applies, the probability of such a large deviation is about 1 in 800,000.

- From the "tip of the red giant branch",

  Wendy Freedman's group: $H_0 = 69.8 \pm 0.6(\text{stat}) \pm 1.6(\text{sys})$ km sec^{-1} Mpc^{-1}

References: Planck Collaboration, Astron. & Astrophys. 641, A6 (2020) [arXiv:1807.06209v4], Table 2, Col. 6

A. Riess et al., Astrophys. J. 934, L7 (2022) [arXiv:2112.04510].

W. Freedman, Astrophys. J. 919, 16 (2021) [arXiv:2106.15656].

The Hubble Diagram: Radiation Flux vs. Redshift

If we live in a universe like we have described, what do we expect to find if we measure the energy flux from a "standard candle" as a function of its redshift?

Consider closed universe:

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \;.$$

We will be interested in tracing radial trajectories, so we can simplify the radial metric by a change of variables

$$\boxed{\sin\psi \equiv \sqrt{k}\,r \;.}$$

Then

$$d\psi = \frac{\sqrt{k}\,dr}{\cos\psi} = \frac{\sqrt{k}\,dr}{\sqrt{1-kr^2}} \;,$$

$$ds^2 = -c^2\,dt^2 + a^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \ .$$

$$d\psi = \frac{\sqrt{k}\,dr}{\cos\psi} = \frac{\sqrt{k}\,dr}{\sqrt{1-kr^2}} \ ,$$

---

and the metric simplifies to

$$\boxed{ds^2 = -c^2\,dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \ ,}$$

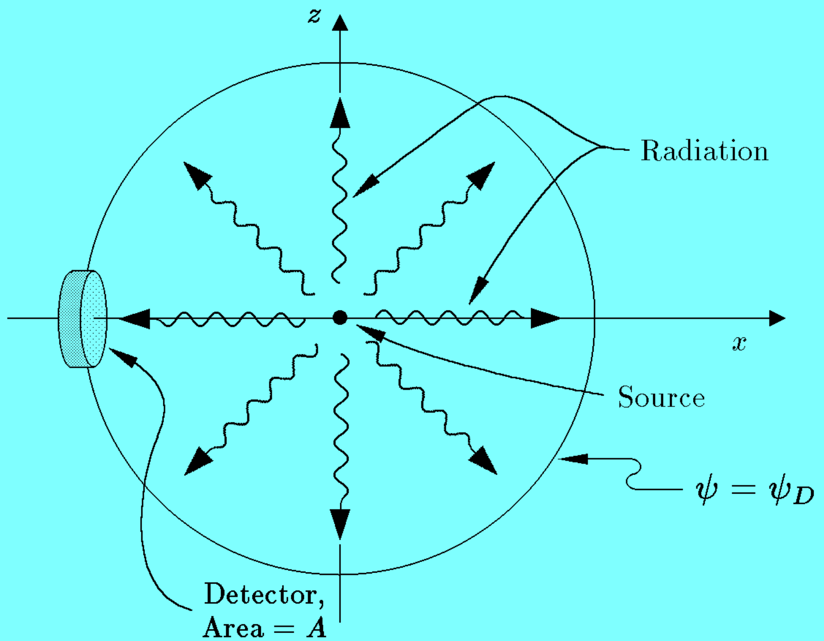
where

$$\tilde{a}(t) \equiv \frac{a(t)}{\sqrt{k}} \ .$$

Note: ψ is in fact the same angle ψ that we used in our construction of the closed-universe metric: it is the angle from the w-axis.

Geometry of Flux Calculation

$$ds^2 = -c^2\, dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right\}$$



The fraction of the photons hitting the sphere that hit the detector is just the ratio of the areas:

$$\text{fraction} = \frac{\text{area of detector}}{\text{area of sphere}} = \frac{A}{4\pi\tilde{a}^2(t_0)\sin^2\psi_D} \ .$$

☆ The power hitting the sphere is less than the power P emitted by the source by two factors of $(1 + z_S)$, where z_S is the redshift of the source: one factor due to redshift of each photon, and one factor due to the redshift of the rate of arrival of photons.

$$\boxed{P_{\text{received}} = \frac{P}{(1 + z_S)^2} \frac{A}{4\pi \tilde{a}^2(t_0) \sin^2 \psi_D} \ .}$$

Flux $J = P_{\text{received}}/A$.

Expressing the Result in Terms of Astronomical Quantities

$$\Omega_{k,0} \equiv -\frac{kc^2}{a^2(t_0)H_0^2} \quad \Longrightarrow \quad \tilde{a}(t_0) = \frac{cH_0^{-1}}{\sqrt{|\Omega_{k,0}|}} \; .$$

But we must still express ψ_D in terms of z_S. Since

$$ds^2 = -c^2\,dt^2 + \tilde{a}^2(t)\left\{d\psi^2 + \sin^2\psi\left(d\theta^2 + \sin^2\theta\,d\phi^2\right)\right\} \;,$$

the equation for a null trajectory is

$$0 = -c^2\,dt^2 + \tilde{a}^2(t)\,d\psi^2 \quad \Longrightarrow \quad \frac{d\psi}{dt} = \frac{c}{\tilde{a}(t)} \; .$$

$$0 = -c^2\,dt^2 + \tilde{a}^2(t)\,d\psi^2 \quad \Longrightarrow \quad \frac{d\psi}{dt} = \frac{c}{\tilde{a}(t)}\ .$$

The first-order Friedmann equation implies

$$H^2 = \left(\frac{\dot{\tilde{a}}}{\tilde{a}}\right)^2 = \frac{H_0^2}{x^4}\left(\Omega_{m,0}\,x + \Omega_{\rm rad,0} + \Omega_{\rm vac,0}\,x^4 + \Omega_{k,0}\,x^2\right)\ ,$$

where

$$x = \frac{a(t)}{a(t_0)} = \frac{\tilde{a}(t)}{\tilde{a}(t_0)}\ .$$

The coordinate distance that the light pulse can travel between t_S (when it left the source) and t_0 (now) is

$$\psi(z_S) = \int_{t_S}^{t_0} \frac{c}{\tilde{a}(t)}\,dt\ .$$

Changing variables to z, with

$$1 + z = \frac{\tilde{a}(t_0)}{\tilde{a}(t)} \ .$$

Then

$$dz = -\frac{\tilde{a}(t_0)}{\tilde{a}(t)^2}\dot{\tilde{a}}(t)\,dt = -\tilde{a}(t_0)H(t)\,\frac{dt}{\tilde{a}(t)} \ .$$

The integration becomes

$$\psi(z_S) = \frac{1}{\tilde{a}(t_0)}\int_0^{z_S} \frac{c}{H(z)}\,dz \ .$$

$$\psi(z_S) = \frac{1}{\tilde{a}(t_0)} \int_0^{z_S} \frac{c}{H(z)}\, dz\ .$$

In this expression we can replace $\tilde{a}(t_0)\ H(z)$ using our previous equations. This gives our final expression for $\psi(z_S)$:

$$\boxed{\begin{aligned}\psi(z_S) = {} & \sqrt{|\Omega_{k,0}|} \\ & \times \int_0^{z_S} \frac{dz}{\sqrt{\Omega_{m,0}(1+z)^3 + \Omega_{\text{rad},0}(1+z)^4 + \Omega_{\text{vac},0} + \Omega_{k,0}(1+z)^2}}\ .\end{aligned}}$$

Using this in our previous expression for J

$$\boxed{J = \frac{P H_0^2 |\Omega_{k,0}|}{4\pi(1+z_S)^2 c^2 \sin^2 \psi(z_S)}\ ,}$$

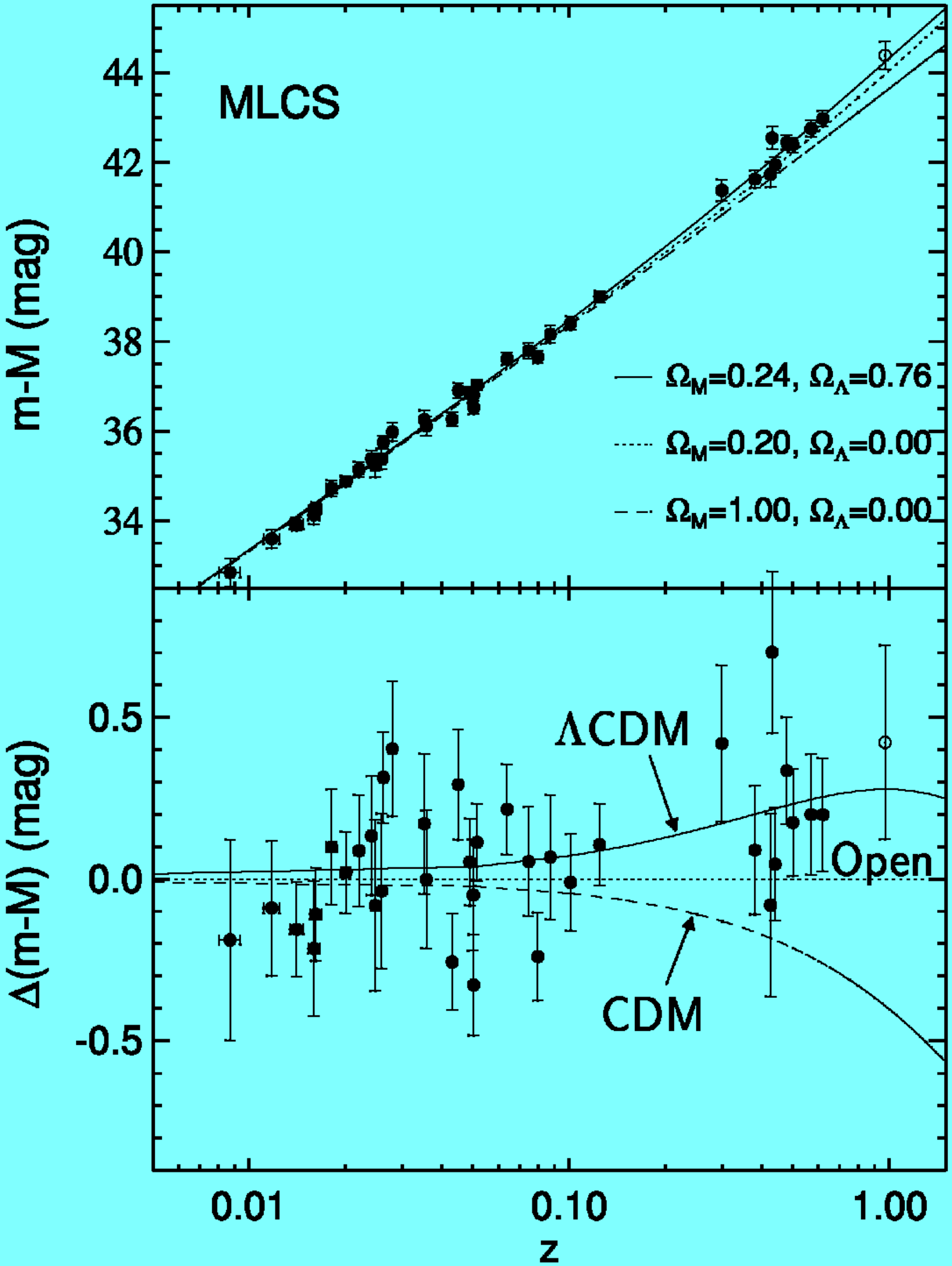
Supernovae Type Ia as Standard Candles

Supernovae Type Ia are believed to be the result of a binary system containing a white dwarf — a stellar remnant that has burned its nuclear fuel, and is supported by electron degeneracy pressure. As the white dwarf accretes gas from its companion star, its mass builds up to 1.4 $M_\odot$, the Chandrasekhar limit, the maximum mass that can be supported by electron degeneracy pressure. The star then collapses, leading to a supernova explosion. Because the Chandrasekhar limit is fixed by physics, all SN Ia are very similar in power output.

There are still some known variations in power output, but they are found to be correlated with the shape of the light curve: if the light curve rises and falls slowly, the supernova is brighter than average.

The properties of SN Ia are known best from observation — theory lags behind.

IF you would like to learn more about this, see Ryden, Section 6.5 (which we skipped — you should not feel obligated to read this).



Hubble diagram from Riess *et al.*, *Astronomical Journal* **116**, No. 3, 1009 (1998) [http://arXiv.org/abs/astro-ph/9805201].

(High-z Supernova Search Team)

Dimmer Supernovae Imply Acceleration

- ☆ The acceleration of the universe is deduced from the fact that distant supernovae appear to be 20-30% dimmer than expected.
- ☆ Why does dimness imply acceleration?
  - Consider a supernova of specified apparent brightness.
  - "Dimmer" implies data point is to the left of where expected — at lower z.
  - Lower z implies slower recession, which implies that the universe was expanding slower than expected in the past — hence, acceleration!

Other Possible Explanations for Dimness

☆ **Absorption by dust.**

- **But absorption usually reddens** the spectrum. This would have to be **"gray" dust,** absorbing uniformly at all observed wavelengths. Such dust is possible, but not known to exist anywhere.
- Dust would most likely be in the host galaxy, which would cause **variable absorption,** depending on SN location in galaxy. Such variability is not seen.

☆ **Chemical evolution** of heavy element abundance.

- But nearby and distant SN Ia look essentially identical.
- For nearby SN Ia, heavy element abundance varies, and does not appear to affect brightness.

☆ Additional evidence against dust or chemical evolution: A SN Ia has been found at $z = 1.7$, which is early enough to be in the decelerating era of the vacuum energy density model. It is consistent with deceleration, but not consistent with either models of absorption or chemical evolution.

Evidence for the Accelerating Universe

1) Supernova Data: distant SN Ia are dimmer than expected by about 20–30%.

2) Cosmic Microwave Background (CMB) anisotropies: gives $\Omega_{\rm vac}$ close to SN value. Also gives $\Omega_{\rm tot} = 1$ to $1/2\%$ accuracy, which cannot be accounted for without dark energy.

3) Inclusion of $\Omega_{\rm vac} \approx 0.70$ makes the age of the universe consistent with the age of the oldest stars.

Evidence for the Accelerating Universe

1) Supernova Data: distant SN Ia are dimmer than expected by about 20–30%.

2) Cosmic Microwave Background (CMB) anisotropies: gives $\Omega_{\rm vac}$ close to SN value. Also gives $\Omega_{\rm tot} = 1$ to 1/2% accuracy, which cannot be accounted for without dark energy.

3) Inclusion of $\Omega_{\rm vac} \approx 0.70$ makes the age of the universe consistent with the age of the oldest stars.

☆ With the 3 arguments together, the case for the accelerating universe and $\Omega_{\rm dark\ energy} \approx 0.70$ has persuaded almost everyone.

Evidence for the Accelerating Universe

1) Supernova Data: distant SN Ia are dimmer than expected by about 20–30%.

2) Cosmic Microwave Background (CMB) anisotropies: gives Ω_{vac} close to SN value. Also gives $\Omega_{\text{tot}} = 1$ to 1/2% accuracy, which cannot be accounted for without dark energy.

3) Inclusion of $\Omega_{\text{vac}} \approx 0.70$ makes the age of the universe consistent with the age of the oldest stars.

☆ With the 3 arguments together, the case for the accelerating universe and $\Omega_{\text{dark energy}} \approx 0.70$ has persuaded almost everyone.

☆ The simplest explanation for dark energy is vacuum energy, but "quintessence" is also possible.

Particle Physics of a Cosmological Constant

☆ $$\boxed{u_{\text{vac}} = \rho_{\text{vac}} c^2 = \frac{\Lambda c^4}{8\pi G}}$$

☆ Contributions to vacuum energy density:

1) Quantum fluctuations of the photon and other bosonic fields: positive and divergent.

2) Quantum fluctuations of the electron and other fermionic fields: negative and divergent.

3) Fields with nonzero values in the vacuum, like the Higgs field.

★ If infinities are cut off at the Planck scale (quantum gravity scale), then infinities become finite, but

☆ If infinities are cut off at the Planck scale (quantum gravity scale), then infinities become finite, but

$>$ 120 orders of magnitude too large!

- ☆ If infinities are cut off at the Planck scale (quantum gravity scale), then infinities become finite, but

  $>$ 120 orders of magnitude too large!

- ☆ For lack of a better explanation, many cosmologists (including Steve Weinberg and yours truly) seriously discuss the possibility that the vacuum energy density is determined by "anthropic" selection effects: that is, maybe there are many types of vacuum (as predicted by string theory), with different vacuum energy densities, with most vacuum energy densities roughly 120 orders of magnitude larger than ours. Maybe we live in a very low energy density vacuum because that is where almost all living beings reside.

- ☆ If infinities are cut off at the Planck scale (quantum gravity scale), then infinities become finite, but

$$\boxed{> \textbf{120 orders of magnitude too large!}}$$

- ☆ For lack of a better explanation, many cosmologists (including Steve Weinberg and yours truly) seriously discuss the possibility that the vacuum energy density is determined by "anthropic" selection effects: that is, maybe there are many types of vacuum (as predicted by string theory), with different vacuum energy densities, with most vacuum energy densities roughly 120 orders of magnitude larger than ours. Maybe we live in a very low energy density vacuum because that is where almost all living beings reside. A large vacuum energy density would cause the universe to rapidly fly apart (if positive) or implode (if negative), so life could not form.

8.286 Class 22
November 23, 2022

PROBLEMS OF THE CONVENTIONAL (NON-INFLATIONARY) HOT BIG BANG MODEL

Particle Physics of a Cosmological Constant

- $$\boxed{u_{\text{vac}} = \rho_{\text{vac}} c^2 = \frac{\Lambda c^4}{8\pi G}}$$

- Contributions to vacuum energy density:

  1) Quantum fluctuations of the photon and other bosonic fields: positive and divergent.

  2) Quantum fluctuations of the electron and other fermionic fields: negative and divergent.

  3) Fields with nonzero values in the vacuum, like the Higgs field.

- ☆ If infinities are cut off at the Planck scale (quantum gravity scale), then infinities become finite, but

$>$ 120 orders of magnitude too large!

- ☆ For lack of a better explanation, many cosmologists (including Steve Weinberg and yours truly) seriously discuss the possibility that the vacuum energy density is determined by "anthropic" selection effects: that is, maybe there are many types of vacuum (as predicted by string theory), with different vacuum energy densities, with most vacuum energy densities roughly 120 orders of magnitude larger than ours. Maybe we live in a very low energy density vacuum because that is where almost all living beings reside. A large vacuum energy density would cause the universe to rapidly fly apart (if positive) or implode (if negative), so life could not form.

Anthropic Selection Effects and the String Theory Landscape

- ☆ Since the inception of string theory, theorists have sought to find the vacuum of string theory — with no success.
- ☆ Since about 2000, most string theorists have come to believe that there is no unique vacuum.
- ☆ Instead, there are perhaps 10^{500} or more long-lived metastable states, any one of which could serve as a substrate for a pocket universe. This is the landscape!
- ☆ Eternal inflation, which we will talk about later, can lead to an infinite number of "pocket universes," of which one would be the universe in which we live. The pocket universes are filled with different types of vacuum, very likely providing examples of every type of vacuum in the string theory landscape.
- ☆ Although string theory would govern everywhere, each type of vacuum would have its own low-energy physics — its own "standard model," its own "constants" of nature, and its own vacuum energy density.

★ If the landscape has 10^{500} vacua, and a fraction 10^{-120} have small vacuum energy densities like our universe, then we expect about

$$10^{-120} \times 10^{500} = 10^{380}$$

vacua with low energy densities like ours.

☆ If the landscape has 10^{500} vacua, and a fraction 10^{-120} have small vacuum energy densities like our universe, then we expect about

$$10^{-120} \times 10^{500} = 10^{380}$$

vacua with low energy densities like ours.

☆ But how could we explain why we are living in such a fantastically unusual type of vacuum?

- If the landscape has 10^{500} vacua, and a fraction 10^{-120} have small vacuum energy densities like our universe, then we expect about

$$10^{-120} \times 10^{500} = 10^{380}$$

vacua with low energy densities like ours.

- But how could we explain why we are living in such a fantastically unusual type of vacuum?

- Possible answer: maybe it is a selection effect. I.e., maybe life only forms where the vacuum energy density is unusually small.

- ☆ As early as 1987, Steve Weinberg pointed out that the vacuum energy density might be explained as a selection effect.
- ☆ Maybe the vacuum energy density ***IS*** huge in most pocket universes. Nonetheless, we need to remember that vacuum energy causes the expansion of the universe to accelerate. If large and negative, the universe quickly implodes. If large and positive, the universe flies apart before galaxies can form. It is plausible, therefore, that life can arise only if the vacuum energy density is very near zero.
- ☆ In 1998 Martel, Shapiro, and Weinberg made a serious calculation of the effect of the vacuum energy density on galaxy formation. They found that to within a factor of order 5, they could "explain" why the vacuum energy density is as small as what we measure.

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- ☆ A number of physicists regard an "anthropic" explanation as simply giving up.

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.
- In my opinion, the selection effect explanation is both logical and scientific.

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.
- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- ☆ A number of physicists regard an "anthropic" explanation as simply giving up.

- ☆ In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe.

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.

- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe. Is this explained by Divine Providence?

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an “anthropic” explanation as simply giving up.
- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe. Is this explained by Divine Providence? By extremely good luck?

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.

- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe. Is this explained by Divine Providence? By extremely good luck? Or is it a selection effect?

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.
- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe. Is this explained by Divine Providence? By extremely good luck? Or is it a selection effect? I think almost all of us would consider it a selection effect.

Anthropic Selection Effects: Does the Selection Effect Explanation Amount to Giving Up?

- A number of physicists regard an "anthropic" explanation as simply giving up.
- In my opinion, the selection effect explanation is both logical and scientific. I think we all accept anthropic explanations for some facts — for example, the fact that we find ourselves living on the surface of a planet, in spite of the fact that planetary surfaces (say ± 1 km) make up an unbelievably small fraction (maybe 10^{-35}) of the volume of the visible universe. Is this explained by Divine Providence? By extremely good luck? Or is it a selection effect? I think almost all of us would consider it a selection effect.
- But I would advocate that anthropic explanations be thought of as the explanation of last resort — the best evidence for an anthropic explanation is the absence of any other.

The Horizon/Homogeneity Problem

The Horizon/Homogeneity Problem

- ☆ General question: how can we explain the large-scale uniformity of the universe?

The Horizon/Homogeneity Problem

- General question: how can we explain the large-scale uniformity of the universe?
- Possible answer: maybe the universe just started out uniform.

The Horizon/Homogeneity Problem

- ☆ General question: how can we explain the large-scale uniformity of the universe?
- ☆ Possible answer: maybe the universe just started out uniform.
  - There is no argument that excludes this possibility, since we don't know how the universe came into being.

The Horizon/Homogeneity Problem

- ☆ General question: how can we explain the large-scale uniformity of the universe?
- ☆ Possible answer: maybe the universe just started out uniform.
  - There is no argument that excludes this possibility, since we don't know how the universe came into being.
  - However, if possible, it seems better to explain the properties of the universe in terms of things that we can understand, rather than to attribute them to things that we don't understand.

The Horizon in Cosmology

- The concept of a horizon was first introduced into cosmology by Wolfgang Rindler in 1956.

- The "horizon problem" was discussed (not by that name) in at least two early textbooks in general relativity and cosmology: Weinberg's *Gravitation and Cosmology* (1972), and Misner, Thorne, and Wheeler's (MTW's) *Gravitation* (1973).

The Cosmic Microwave Background

☆ The strongest evidence for the uniformity of the universe comes from the CMB, since it has been measured so precisely.

The Cosmic Microwave Background

- ☆ The strongest evidence for the uniformity of the universe comes from the CMB, since it has been measured so precisely.
- ☆ The radiation appears slightly hotter in one direction than in the opposite direction, by about one part in a thousand — but this nonuniformity can be attributed to our motion through the background radiation.
- ☆ Once this effect is subtracted out, using best-fit parameters for the velocity, it is found that the residual temperature pattern is uniform to a few parts in 10^5.

The Cosmic Microwave Background

- ☆ The strongest evidence for the uniformity of the universe comes from the CMB, since it has been measured so precisely.
- ☆ The radiation appears slightly hotter in one direction than in the opposite direction, by about one part in a thousand — but this nonuniformity can be attributed to our motion through the background radiation.
- ☆ Once this effect is subtracted out, using best-fit parameters for the velocity, it is found that the residual temperature pattern is uniform to a few parts in 10^5.
- ☆ Could this be simply the phenomenon of thermal equilibrium? If you put an ice cube on the sidewalk on a hot summer day, it melts and comes to the same temperature as the sidewalk.

The Cosmic Microwave Background

- ☆ The strongest evidence for the uniformity of the universe comes from the CMB, since it has been measured so precisely.
- ☆ The radiation appears slightly hotter in one direction than in the opposite direction, by about one part in a thousand — but this nonuniformity can be attributed to our motion through the background radiation.
- ☆ Once this effect is subtracted out, using best-fit parameters for the velocity, it is found that the residual temperature pattern is uniform to a few parts in 10^5.
- ☆ Could this be simply the phenomenon of thermal equilibrium? If you put an ice cube on the sidewalk on a hot summer day, it melts and comes to the same temperature as the sidewalk.
- ☆ **BUT: in the conventional model of the universe, it did not have enough time for thermal equilibrium to explain the uniformity, if we assume that it did not start out uniform. If no matter, energy, or information can travel faster than light, then it is simply not possible.**

Basic History of the CMB

- In conventional cosmological model, the universe at the earliest times was radiation-dominated. It started to be matter-dominated at $t_{\rm eq} \approx 50,000$ years, the time of matter-radiation equality.
- At the time of *decoupling* $t_d \approx 380,000$ years, the universe cooled to about 3000 K, by which time the hydrogen (and some helium) combined so thoroughly that free electrons were very rare. At earlier times, the universe was in a mainly plasma phase, with many free electrons, and photons were essentially frozen with the matter. At later times, the universe was transparent, so photons have traveled on straight lines. We can say that the CMB was released at about 380,000 years.
- Since the photons have been mainly traveling on straight lines since $t = t_d$, they have all traveled the same distance. Therefore the locations from which they were released form a sphere centered on us. This sphere is called the *surface of last scattering*, since the photons that we receive now in the CMB was mostly scattered for the last time on or very near this surface.

- As we learned in Lecture Notes 4, the horizon distance is defined as the present distance of the furthest particles from which light has had time to reach us, since the beginning of the universe.
- For a matter-dominated flat universe, the horizon distance at time t is $3ct$, while for a radiation-dominated universe, it is $2ct$.
- At $t = t_d$ the universe was well into the matter-dominated phase, so we can approximate the horizon distance as

$$\boxed{\ell_h(t_d) \approx 3ct_d \approx 1,100,000 \text{ light-years.}}$$

For comparison, we would like to calculate the radius of the surface of last scattering at time t_d, since this region is the origin of the photons that we are now receiving in the CMB. I will denote the physical radius of the surface of last scattering, at time t, as $\ell_p(t)$.

$$\boxed{\ell_h(t_d) \approx 3ct_d \approx 1,100,000 \text{ light-years.}}$$

For comparison, we would like to calculate the radius of the surface of last scattering at time t_d, since this region is the origin of the photons that we are now receiving in the CMB. I will denote the physical radius of the surface of last scattering, at time t, as $\ell_p(t)$.

- ☆ To calculate $\ell_p(t_d)$, I will make the crude approximation that the universe has been matter-dominated at all times. (We will find that this *horizon problem* is very severe, so even if our calculation is wrong by a factor of 2, it won't matter.)
- ☆ Strategy: find $\ell_p(t_0)$, and scale to find $\ell_p(t_d)$. Under the assumption of a flat matter-dominated universe, we learned that the physical distance today to an object at redshift z is

$$\ell_p(t_0) = 2cH_0^{-1}\left[1 - \frac{1}{\sqrt{1+z}}\right] .$$

$$\ell_p(t_0) = 2cH_0^{-1}\left[1 - \frac{1}{\sqrt{1+z}}\right] \ .$$

- The redshift of the surface of last scattering is about

$$1 + z = \frac{a(t_0)}{a(t_d)} = \frac{3000\ \mathrm{K}}{2.7\ \mathrm{K}} \approx 1100 \ .$$

- If we take $H_0 = 67.7$ km-s^{-1}-Mpc^{-1}, one finds that $H_0^{-1} \approx 14.4 \times 10^9$ yr and $\ell_p(t_0) \approx 28.0 \times 10^9$ light-yr. (Note that $\ell_p(t_0)$ is equal to 0.970 times the current horizon distance — very close.)

- To find $\ell_p(t_d)$, just use the fact that the redshift is related to the scale factor:

$$\boxed{\begin{aligned} \ell_p(t_d) &= \frac{a(t_d)}{a(t_0)}\ell_p(t_0) \\ &\approx \frac{1}{1100} \times 28.0 \times 10^9\ \text{lt-yr} \approx 2.55 \times 10^7\ \text{lt-yr} \ . \end{aligned}}$$

$$\boxed{\ell_h(t_d) \approx 3ct_d \approx 1,100,000 \text{ light-years.}}$$

$$\begin{aligned} \ell_p(t_d) &= \frac{a(t_d)}{a(t_0)}\ell_p(t_0) \\ &\approx \frac{1}{1100} \times 28.0 \times 10^9 \text{ lt-yr} \approx 2.55 \times 10^7 \text{ lt-yr} \quad . \end{aligned}$$

☆ Comparison: At the time of decoupling, the ratio of the radius of the surface of last scattering to the horizon distance was

$$\boxed{\frac{\ell_p(t_d)}{\ell_h(t_d)} \approx \frac{2.55 \times 10^7 \text{ lt-yr}}{1.1 \times 10^6 \text{ lt-yr}} \approx 23 \ .}$$

Summary of the Horizon Problem

Suppose that one detects the cosmic microwave background in a certain direction in the sky, and suppose that one also detects the radiation from precisely the opposite direction. At the time of emission, the sources of these two signals were separated from each other by about 46 horizon distances. Thus it is absolutely impossible, within the context of this model, for these two sources to have come into thermal equilibrium by any physical process.

Although our calculation ignored the dark energy phase, we have found in previous examples that such calculations are wrong by some tens of a percent. (For example we found $t_{\rm eq} \approx 75,000$ years, when it should have been about 50,000 years.) Since $46 \gg 1$, there is no way that a more accurate calculation could cause this problem to go away.

The Flatness Problem

- A second problem of the conventional cosmological model is the *flatness problem*: why was the value of Ω in the early universe so extraordinarily close to 1?

- Today we know, according to the Planck satellite team analysis (2018), that

$$\boxed{\Omega_0 = 0.9993 \pm 0.0037}$$

at 95% confidence. I.e., $\Omega = 1$ to better than 1/2 of 1%.

- As we will see, this implies that Ω in the early universe was extaordinarily close to 1. For example, at $t = 1$ second,

$$\boxed{|\Omega - 1|_{t=1\text{ sec}} < 10^{-18}\ .}$$

- The underlying fact is that the value $\Omega = 1$ is a point of unstable equilibrium, something like a pencil balancing on its point. If Ω is ever ***exactly*** equal to one, it will remain equal to one forever — that is, a flat $(k = 0)$ universe remains flat. However, if Ω is ever slightly larger than one, it will rapidly grow toward infinity; if Ω is ever slightly smaller than one, it will rapidly fall toward zero. For Ω to be anywhere near 1 today, Ω in the early universe must have been extraordinarily close to one.
- Like the horizon problem, the flatness problem could in principle be solved by the initial conditions of the universe: maybe the universe began with $\Omega \equiv 1$.
    - But, like the horizon problem, it seems better to explain the properties of the universe, if we can, in terms of things that we can understand, rather than to attribute them to things that we don't understand.

History of the Flatness Problem

The mathematics behind the flatness problem was undoubtedly known to almost anyone who has worked on the big bang theory from the 1920's onward, but apparently the first people to consider it a problem in the sense described here were Robert Dicke and P.J.E. Peebles, who published a discussion in 1979.*

*R.H. Dicke and P.J.E. Peebles, "The big bang cosmology — enigmas and nostrums," in **General Relativity: An Einstein Centenary Survey**, eds: S.W. Hawking and W. Israel, Cambridge University Press (1979).

The Mathematics of the Flatness Problem

Start with the first-order Friedmann equation:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} \; .$$

Remembering that $\Omega = \rho/\rho_c$ and that $\rho_c = 3H^2/(8\pi G)$, one can divide both sides of the equation by H^2 to find

$$1 = \frac{\rho}{\rho_c} - \frac{kc^2}{a^2 H^2} \quad \Longrightarrow \quad \boxed{\Omega - 1 = \frac{kc^2}{a^2 H^2} \; .}$$

Evolution of $\Omega - 1$ During the Radiation-Dominated Phase

$$\Omega - 1 = \frac{kc^2}{a^2 H^2} \ .$$

For a (nearly) flat radiation-dominated universe, $a(t) \propto t^{1/2}$, so $H = \dot{a}/a = 1/(2t)$. So

$$\Omega - 1 \propto \left(\frac{1}{t^{1/2}}\right)^2 \left(\frac{1}{t^{-1}}\right)^2 \propto t \quad \text{(radiation dominated).}$$

Evolution of $\Omega - 1$ During the Matter-Dominated Phase

$$\boxed{\Omega - 1 = \frac{kc^2}{a^2 H^2}\ .}$$

For a (nearly) flat matter-dominated universe, $a(t) \propto t^{2/3}$, so $H = \dot{a}/a = 2/(3t)$. So

$$\boxed{\Omega - 1 \propto \left(\frac{1}{t^{2/3}}\right)^2 \left(\frac{1}{t^{-1}}\right)^2 \propto t^{2/3} \quad \text{(matter-dominated).}}$$

Tracing $\Omega - 1$ from Now to 1 Second

Today,

$$|\Omega_0 - 1| < .01 \ .$$

I will do a crude calculation, treating the universe as matter dominated from 50,000 years to the present, and as radiation-dominated from 1 second to 50,000 years.

During the matter-dominated phase,

$$(\Omega - 1)_{t=50{,}000 \text{ yr}} \approx \left(\frac{50{,}000}{13.8 \times 10^9}\right)^{2/3} (\Omega_0 - 1) \approx 2.36 \times 10^{-4} \, (\Omega_0 - 1) \ .$$

$$|\Omega_0 - 1| < .01 \ .$$

$$(\Omega - 1)_{t=50{,}000 \text{ yr}} \approx \left(\frac{50{,}000}{13.8 \times 10^9} \right)^{2/3} (\Omega_0 - 1) \approx 2.36 \times 10^{-4} \, (\Omega_0 - 1) \ .$$

During the radiation-dominated phase,

$$\begin{aligned} (\Omega - 1)_{t=1 \text{ sec}} &\approx \left(\frac{1 \text{ sec}}{50{,}000 \text{ yr}} \right) (\Omega - 1)_{t=50{,}000 \text{ yr}} \\ &\approx 1.49 \times 10^{-16} \, (\Omega_0 - 1) \ . \end{aligned}$$

The conclusion is therefore

$$|\Omega - 1|_{t=1 \text{ sec}} < 10^{-18} \ .$$

The conclusion is therefore

$$\boxed{|\Omega - 1|_{t=1\text{ sec}} < 10^{-18}\ .}$$

Even if we put ourselves mentally back into 1979, we would have said that $0.1 < \Omega_0 < 2$, so $|\Omega_0 - 1| < 1$, and would have concluded that

$$\boxed{|\Omega - 1|_{t=1\text{ sec}} < 10^{-16}\ .}$$

The Dicke & Peebles paper, that first pointed out this problem, also considered $t = 1$ second, but concluded (without showing the details) that

$$\boxed{|\Omega - 1|_{t=1\text{ sec}} < 10^{-14}\ .}$$

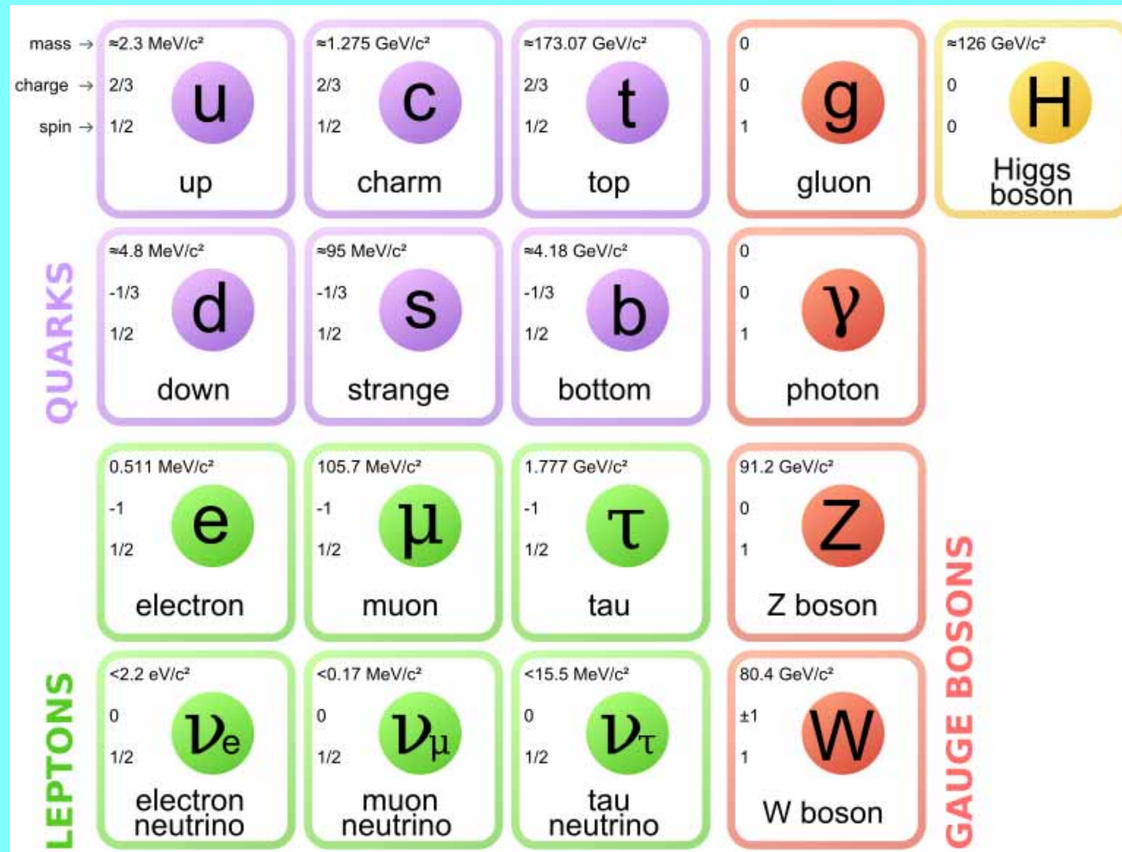
They were perhaps more conservative, but concluded nonetheless that this extreme fine-tuning cried out for an explanation.

8.286 Class 23
November 28, 2022

GRAND UNIFIED THEORIES
AND THE
MAGNETIC MONOPOLE PROBLEM

The Standard Model of Particle Physics

Particle Content:



Wikimedia Commons. Source: PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group.

Quarks are Colored

- A quark is specified by its *flavor* [u(p), d(own), c(harmed), s(trange), t(op), b(ottom)], its spin [up or down, along any chosen z axis], whether it is a quark or antiquark, AND ITS COLOR [three choices, often red, blue, or green].
- Quarks that differ only in color are completely indistinguishable, but the color is relevant for the Pauli exclusion principle: one can't have 3 identical quarks all in the lowest energy state, but one can have one red quark, one blue quark, and one green quark.
- Color is also relevant for the way quarks interact. The colors act like a generalized form of electric charge. Two red quarks interact with each other exactly the same way as two blue quarks, but a red quark and a blue quark interact with each other differently.
- Any isolated system of quarks must be a "color singlet". The simplest color singlets — and the only ones known to exist in nature — are 3-quark states (baryons), with equal parts of red, blue, and green, and quark-antiquark states (mesons), with equal parts of each color and its anticolor. If one tries to pull a quark from a proton, it is pulled back with a force independent of distance, as if it were attached by a string. (Called "confinement".)

Gauge Theories: Electromagnetic Example

Fields and potentials*: $\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$, $\vec{B} = \vec{\nabla} \times \vec{A}$.

Four-vector notation: $A_\mu = \left(-\frac{\phi}{c}, A_i\right)$, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$,

$E_i = cF_{i,0}$, $B_i = \frac{1}{2}\epsilon_{ijk}F_{jk}$.

Gauge transformations:

$$\phi'(t,\vec{x}) = \phi(t,\vec{x}) - \frac{\partial \Lambda(t,\vec{x})}{\partial t} \ , \quad \vec{A}'(t,\vec{x}) = \vec{A}(t,\vec{x}) + \vec{\nabla}\Lambda(t,\vec{x}) \ ,$$

or in four-vector notation,

$$A'_\mu(x) = A_\mu(x) + \frac{\partial \Lambda}{\partial x^\mu} \ , \text{ where } x^\mu \equiv (ct, \vec{x}).$$

$\vec{E}$ and $\vec{B}$ are gauge-invariant (i.e., are unchanged by a gauge transformation):

$$\vec{B}' = \vec{\nabla} \times \vec{A}' = \vec{\nabla} \times (\vec{A} + \vec{\nabla}\Lambda) = \vec{\nabla} \times \vec{A} = \vec{B} \ ,$$

*Using the conventions of D.J. Griffiths, *Introduction to Electrodynamics*, Fourth Edition.

$$\phi'(t,\vec{x}) = \phi(t,\vec{x}) - \frac{\partial\Lambda(t,\vec{x})}{\partial t}\ , \quad \vec{A}'(t,\vec{x}) = \vec{A}(t,\vec{x}) + \vec{\nabla}\Lambda(t,\vec{x})\ ,$$

$\vec{E}$ and $\vec{B}$ are gauge-invariant (i.e., are unchanged by a gauge transformation):

$$\vec{B}' = \vec{\nabla}\times\vec{A}' = \vec{\nabla}\times(\vec{A}+\vec{\nabla}\Lambda) = \vec{\nabla}\times\vec{A} = \vec{B}\ ,$$

$$\vec{E}' = -\vec{\nabla}\phi' - \frac{\partial\vec{A}'}{\partial t} = -\vec{\nabla}\left(\phi - \frac{\partial\Lambda}{\partial t}\right) - \frac{\partial}{\partial t}\left(\vec{A}+\vec{\nabla}\Lambda\right)$$

$$= -\vec{\nabla}\phi - \frac{\partial\vec{A}}{\partial t} = \vec{E}\ ,$$

where we used $\vec{\nabla}\times\vec{\nabla}\Lambda \equiv 0$ and $\vec{\nabla}\left(\frac{\partial\Lambda}{\partial t}\right) = \frac{\partial}{\partial t}\vec{\nabla}\Lambda$. So A_μ and A'_μ both satisfy the equations of motion, and describe the SAME physical situation.

Gauge transformations can be combined, forming a group:

$$\Lambda_3(x) = \Lambda_1(x) + \Lambda_2(x)\ .$$

Gauge symmetries are also called local symmetries, since the gauge function $\Lambda(x)$ is an arbitrary function of position and time.

Electromagnetism as a $U(1)$ Gauge Theory

$\Lambda(x)$ is an element of the real numbers.

But if we included an electron field $\psi(x)$, it would transform as

$$\psi'(x) = e^{ie_0\Lambda(x)}\psi(x) \ ,$$

where e_0 is the charge of a proton and $e = 2.71828\ldots$. So we might think of $u(x) \equiv e^{ie_0\Lambda(x)}$ as describing the gauge transformation. u contains LESS information than Λ, since it defines Λ only mod $2\pi/e_0$.

But u is enough to define the gauge transformation, since

$$\frac{\partial \Lambda}{\partial x^\mu} = \frac{1}{ie_0}\, e^{-ie_0\Lambda(x)}\, \frac{\partial}{\partial x^\mu}\, e^{ie_0\Lambda(x)} \ .$$

u is an element of the group $U(1)$, the group of complex phases $u = e^{i\chi}$, where χ is real. So E&M is a $U(1)$ gauge theory.

Gauge Groups of the Standard Model

$U(1)$ is abelian (commutative), but Yang and Mills showed in 1954 how to construct a nonabelian gauge theory. The standard model contains the following gauge symmetries:

***SU*(3)**: This is the group of 3×3 complex matrices that are

S $\equiv$ Special: they have determinant 1.

U $\equiv$ Unitary: they obey $u^\dagger u = 1$, which means that when they multiply a 1×3 column vector v, they preserve the norm $|v| \equiv \sqrt{v_i^* v_i}$.

***SU*(2)**: The group of 2×2 complex matrices that are special (S) and unitary (U). As you may have learned in quantum mechanics, $SU(2)$ is almost the same as the rotation group in 3D, with a 2:1 group-preserving mapping between $SU(2)$ and the rotation group.

***U*(1)**: The group of complex phases. The $U(1)$ of the standard model is not the $U(1)$ of E&M; instead $U(1)_{\text{E\&M}}$ is a linear combination of the $U(1)$ of the standard model and a rotation about one fixed direction in $SU(2)$.

Combining the groups: the gauge symmetry group of the standard model is usually described as $SU(3)\times SU(2)\times U(1)$. An element of this group is an ordered triplet (u_3, u_2, u_1), where $u_3 \in SU(3)$, $u_2 \in SU(2)$, and $u_1 \in U(1)$, so $SU(3)\times SU(2)\times U(1)$ is really no different from thinking of the 3 symmetries separately.

$SU(3)$ describes the strong interactions, and $SU(2) \times U(1)$ together describe the electromagnetic and weak interactions in a unified way, called the *electroweak interactions*.

$SU(3)$ acts on the quark fields by rotating the 3 "colors" into each other. Thus the strong interactions of the quarks are entirely due to their "colors", which act like charges.

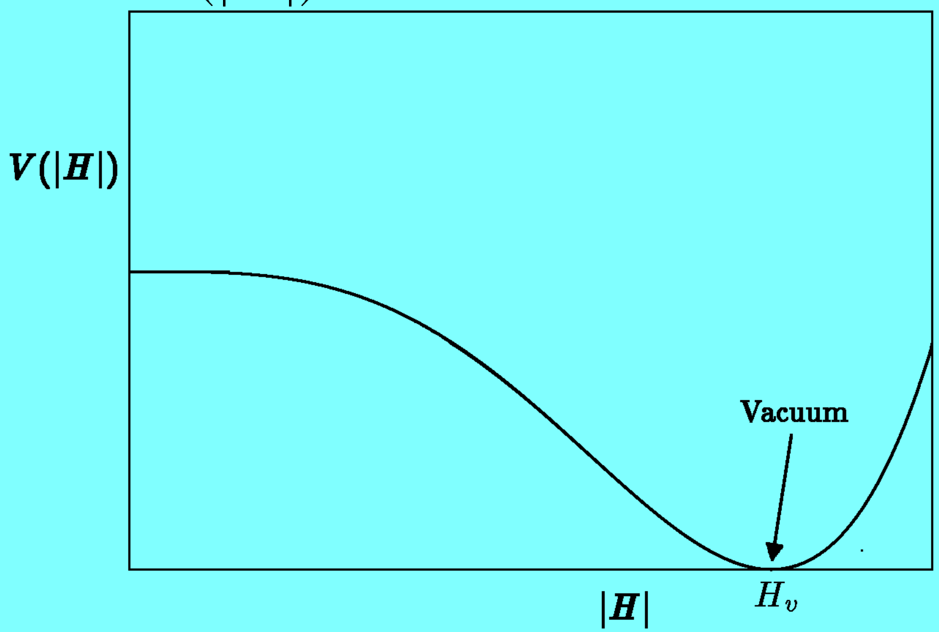
The Higgs Field and Spontaneous Symmetry Breaking

The Higgs field is a complex doublet:

$$H(x) \equiv \begin{pmatrix} h_1(x) \\ h_2(x) \end{pmatrix} \ .$$

Under $SU(2)$ transformations, $H'(x) = u_2(x)H(x)$, where $u_2(x)$ is the complex 2×2 matrix that defines the $SU(2)$ gauge transformation at x. Since the gauge symmetry implies that the potential energy density of the Higgs field $V(H)$ must be gauge-invariant, V can depend only on $|H| \equiv \sqrt{|h_1|^2 + |h_2|^2}$, which is unchanged by $SU(2)$ transformations.

Potential energy function $V(|H|)$:



The minimum is not at $|H| = 0$, but instead at $|H| = H_v$.

$|H| = 0$ is $SU(2)$ gauge-invariant, but $|H| = H_v$ is not. H randomly picks out some direction in the space of 2D complex vectors.

Spontaneous Symmetry Breaking: Whenever the ground state of a system has less symmetry than the underlying laws, it is called spontaneous symmetry breaking. Examples: crystals, ferromagnetism.

Higgs Fields Give Mass to Other Particles

When $H = 0$, all the fundamental particles of the standard model are massless. Furthermore, there is no distinction between the electron e and the electron neutrino ν_e, or between μ and ν_μ, or between τ and ν_τ. (Protons, however, would not be massless — intuitively, most of the proton mass comes from the gluon field that binds the quarks.)

For $|H| \neq 0$, H randomly picks out a direction in the space of 2D complex vectors. Since all directions are otherwise equivalent, we can assume that in the vacuum,

$$H = \begin{pmatrix} H_v \\ 0 \end{pmatrix} .$$

Components of other fields that interact with $\mathrm{Re}\,(h_1)$ then start to behave differently from fields that interact with other components of H.

Mass: mc^2 of a particle is the state of lowest energy above the ground state. In a field theory, this corresponds to a homogeneous oscillation of the field, which in turn corresponds to a particle with zero momentum.

In the free field limit, the field acts exactly like a harmonic oscillator. The first excited state has energy $h\nu = \hbar\omega$ above the ground state. So, $mc^2 = \hbar\omega$.

ω is determined by inertia and the restoring force. When $H = 0$, the standard model interactions provide no restoring forces. Any such restoring force would break gauge invariance.

When $H = \begin{pmatrix} H_v \\ 0 \end{pmatrix}$, the interactions with H creates a restoring force for some components of other fields, giving them a mass. This "Higgs mechanism" creates the distinction between electrons and neutrinos — the electrons are the particles that get a mass, and the neutrinos do not. (Neutrinos are exactly massless in the Standard Model of Particle Physics. There are various ways to modify the model to account for neutrino masses.)

8.286 Class 24
November 30, 2022

GRAND UNIFIED THEORIES
AND THE
MAGNETIC MONOPOLE PROBLEM, PART 2

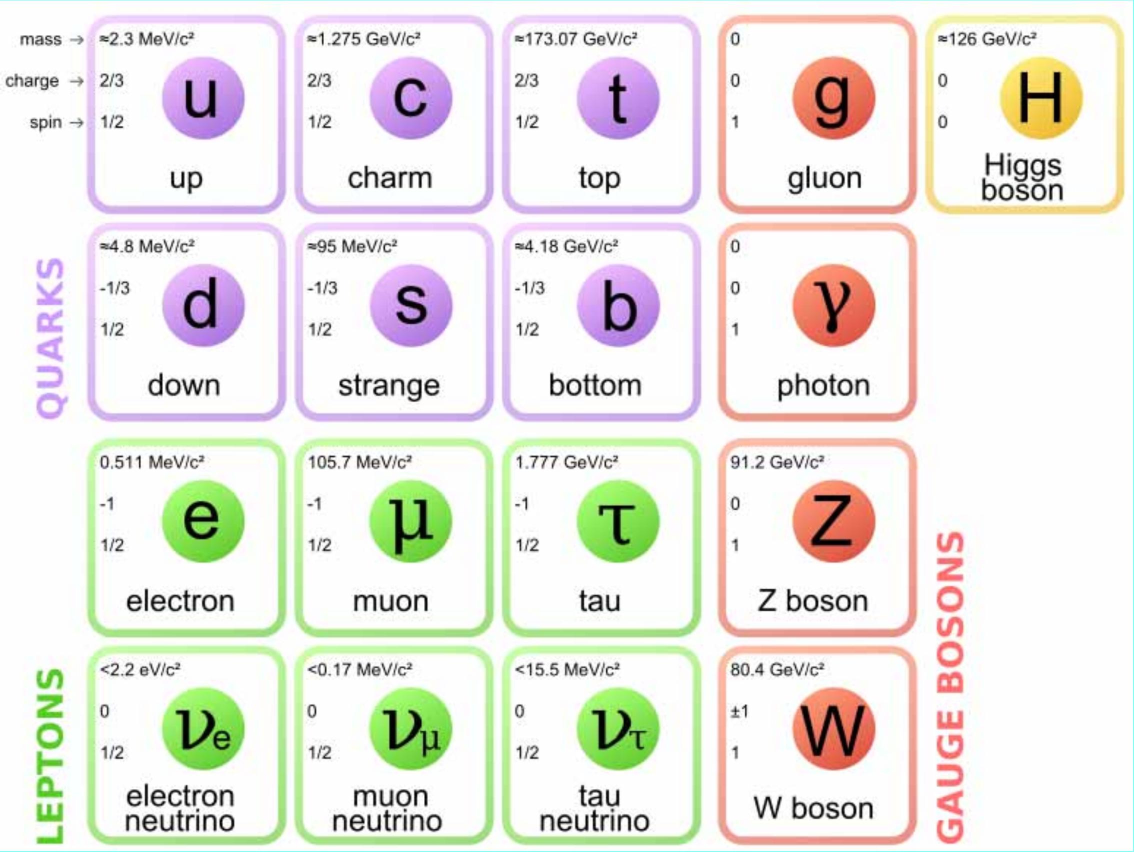
Announcements

★ Quiz 3: Next Wednesday.

Review from last class:

The Standard Model of Particle Physics

Particle Content:



Wikimedia Commons. Source: PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group.

Quarks are Colored

- A quark is specified by its *flavor* [u(p), d(own), c(harmed), s(trange), t(op), b(ottom)], its spin [up or down, along any chosen z axis], whether it is a quark or antiquark, AND ITS COLOR [three choices, often red, blue, or green].
- Color symmetry is exact.
- Colors plays the role of charges, describing how the quarks interact.
- Any isolated state must be a "color singlet".

Gauge Theories: Electromagnetic Example

Fields and potentials: $\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}$, $\vec{B} = \vec{\nabla} \times \vec{A}$.

Four-vector notation: $A_\mu = \left(-\frac{\phi}{c}, A_i\right)$, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$,

$E_i = cF_{i,0}$, $B_i = \frac{1}{2}\epsilon_{ijk}F_{jk}$.

Gauge transformations, in four-vector notation:

$$A'_\mu(x) = A_\mu(x) + \frac{\partial \Lambda}{\partial x^\mu} \ , \ \text{where } x \equiv (ct, \vec{x}).$$

Field configurations $A_\mu(x)$ that are related by a gauge transformation represent the SAME physical situation.

Electromagnetism as a $U(1)$ Gauge Theory

$\Lambda(x)$ is an element of the real numbers.

To construct the gauge transformation, it is sufficient to know

$$u \equiv e^{ie_0 \Lambda(x)} \ ,$$

where e_0 is the charge of a proton and $e = 2.71828\ldots$.

This is LESS information, since we only have to know $\Lambda(x)$ modulo $2\pi/e_0$.

u is an element of the group $U(1)$, the group of complex phases $u = e^{i\chi}$, where χ is real. So E&M is a $U(1)$ gauge theory.

Review from last class:

Gauge Groups of the Standard Model

$U(1)$ is abelian (commutative), but Yang and Mills showed in 1954 how to construct a nonabelian gauge theory. The standard model contains the following gauge symmetries:

***SU*(3)**: This is the group of 3×3 complex matrices that are

S $\equiv$ Special: they have determinant 1.

U $\equiv$ Unitary: they obey $u^\dagger u = 1$, which means that when they multiply a 1×3 column vector v, they preserve the norm $|v| \equiv \sqrt{v_i^* v_i}$.

***SU*(2)**: The group of 2×2 complex matrices that are special (S) and unitary (U). As you may have learned in quantum mechanics, $SU(2)$ is almost the same as the rotation group in 3D, with a 2:1 group-preserving mapping between $SU(2)$ and the rotation group.

***U*(1)**: The group of complex phases. The $U(1)$ of the standard model is not the $U(1)$ of E&M; instead $U(1)_{\mathrm{E\&M}}$ is a linear combination of the $U(1)$ of the standard model and a rotation about one fixed direction in $SU(2)$.

Review from last class:

Combining the groups: the gauge symmetry group of the standard model is usually described as $SU(3)\times SU(2)\times U(1)$. An element of this group is an ordered triplet (u_3, u_2, u_1), where $u_3 \in SU(3)$, $u_2 \in SU(2)$, and $u_1 \in U(1)$, so **$SU(3)\times SU(2)\times U(1)$ is really no different from thinking of the 3 symmetries separately.**

$SU(3)$ describes the strong interactions, and $SU(2) \times U(1)$ together describe the electromagnetic and weak interactions in a unified way, called the *electroweak interactions*.

Gauge theories always have one gauge boson for each parameter of the gauge group:

$SU(3)$: 8 parameters $\Longrightarrow$ 8 gluons.

$SU(2)\times U(1)$: 3 + 1 parameters $\Longrightarrow$ photon and W^+, W^-, and Z.

The gauge symmetry dictates how these particles interact. If the gauge symmetry is not spontaneously broken (to be discussed shortly), the gauge boson is massless, like the photon.

Review from last class:

The Higgs Field and Spontaneous Symmetry Breaking

The Higgs field is a complex doublet:

$$H(x) \equiv \begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} \ .$$

Under $SU(2)$ transformations, $H'(x) = u_2(x)H(x)$, where $u_2(x)$ is the complex 2×2 matrix that defines the $SU(2)$ gauge transformation at x.

Review from last class, but adding $\vec{\phi}$:

The Higgs field is a complex doublet:

$$H(x) \equiv \begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} .$$

Under $SU(2)$ transformations, $H'(x) = u_2(x)H(x)$, where $u_2(x)$ is the complex 2×2 matrix that defines the $SU(2)$ gauge transformation at x.

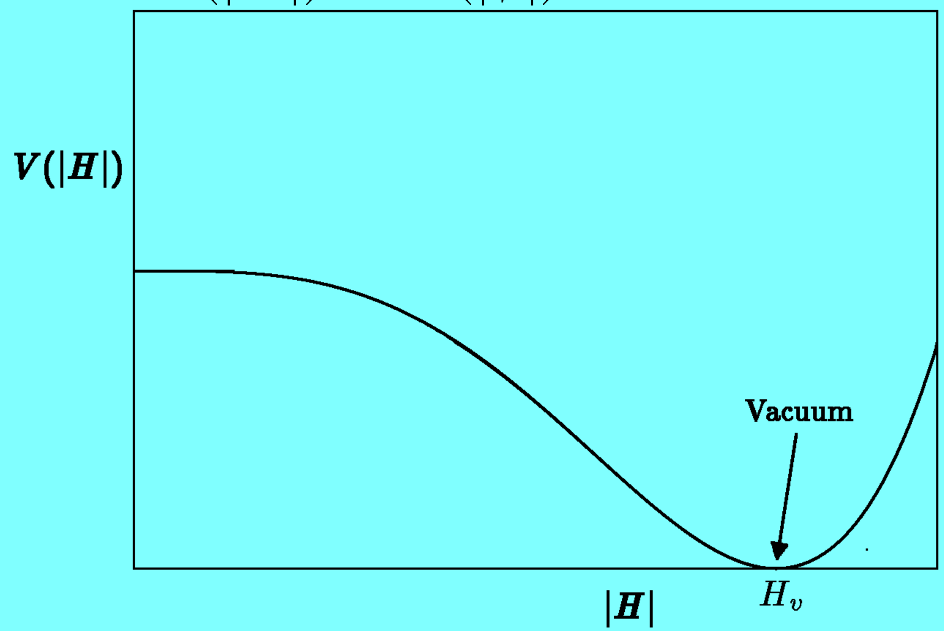
Toy Theory (easier to understand): Consider a "vector Higgs" $\vec{\phi}(x)$, with 3 real components:

$$\vec{\phi}(x) \equiv \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix} .$$

Recall that $SU(2)$ is closely related to the 3D rotation group: there are 2 elements of $SU(2)$ for every element of the rotation group. $\vec{\phi}$ transforms just like any vector under these rotations.

Since the gauge symmetry implies that the potential energy density of the Higgs field $V(H)$ must be gauge-invariant, V can depend only on $|H| \equiv \sqrt{|H_1|^2 + |H_2|^2}$, or in the toy theory, $|\vec{\phi}| \equiv \sqrt{\phi_1^2 + \phi_2^2 + \phi_3^2}$, which is unchanged by $SU(2)$ transformations.

Potential energy function $V(|H|)$ or $V(|\vec{\phi}|)$:



The minimum is not at $|H| = 0$, but instead at $|H| = H_v$.

$|H| = 0$ is $SU(2)$ gauge-invariant, but $|H| = H_v$ is not. H randomly picks out some direction in the space of 2D complex vectors.

In the toy vector Higgs theory, $\vec{\phi} = 0$ is rotationally invariant, but $\vec{\phi} \neq 0$ must pick out some direction. $\vec{\phi}$ is invariant under rotations about $\vec{\phi}$, but not under other rotations. So the vector Higgs "breaks" the 3D rotation group symmetry down to 1D rotations (which is the same as $U(1)$).

Spontaneous Symmetry Breaking

Definition: Whenever the ground state of a system has less symmetry than the underlying laws, it is called spontaneous symmetry breaking. Other examples: crystals, ferromagnetism.

Higgs Fields Give Mass to Other Particles

When $H = 0$, all the fundamental particles of the standard model are massless. Furthermore, there is no distinction between the electron e and the electron neutrino ν_e, or between μ and ν_μ, or between τ and ν_τ. (Protons, however, would not be massless — intuitively, most of the proton mass comes from the gluon field that binds the quarks.)

To describe how H gives mass to the other particles, consider the toy vector Higgs theory. For $|\vec{\phi}| \neq 0$, $\vec{\phi}$ randomly picks out a direction in the 3D space of (ϕ_1, ϕ_2, ϕ_3) . Since all directions are otherwise equivalent, we can assume that in the vacuum,

$$\vec{\phi} = \begin{pmatrix} 0 \\ 0 \\ \phi_v \end{pmatrix} .$$

Components of other fields that interact with ϕ_3 then start to behave differently from fields that interact with other components of $\vec{\phi}$.

Mass: mc^2 of a particle is the state of lowest energy above the ground state. In a field theory, this corresponds to a homogeneous oscillation of the field, which in turn corresponds to a particle with zero momentum.

If we ignore the interactions between fields, the field acts exactly like a harmonic oscillator. The first excited state has energy $h\nu = \hbar\omega$ above the ground state. So, $mc^2 = \hbar\omega$.

ω is determined by inertia and the restoring force. When $\vec{\phi} = 0$, the standard model interactions provide no restoring forces. Any such restoring force would break gauge invariance.

When $\vec{\phi} = \begin{pmatrix} 0 \\ 0 \\ \phi_v \end{pmatrix}$, the interactions with $\vec{\phi}$ create a restoring force for some components of other fields, giving them a mass. (That is, the energy density can contain terms such as $\phi_3\psi^2$, creating a restoring force for the field ψ.) This "Higgs mechanism" creates the distinction between electrons and neutrinos — the electrons are the particles that get a mass, and the neutrinos do not. (Neutrinos are exactly massless in the Standard Model of Particle Physics. There are various ways to modify the model to account for neutrino masses.)

The Higgs mechanism, through the nonzero components of $\vec{\phi}$, also gives a mass to some of the gauge bosons. The gauge bosons that correspond to broken symmetries are given a mass, while the gauge bosons that correspond to unbroken symmetries remain massless.

Beyond the Standard Model

With neutrino masses added, the standard model is spectacularly successful: it agrees with all reliable particle experiments.

Nonetheless, most physicists regard it as incomplete, for at least two types of reasons:

1) It does not include gravity, and it does not include any particle to account for the dark matter. (Maybe black holes can be the dark matter, but that requires a mass distribution that we cannot explain.)

2) The theory appears too inelegant to be the final theory. It contains more arbitrary features and free parameters than one would hope for in a final theory. Why $SU(3)\times SU(2)\times U(1)$? Why three generations of fermions? The original theory had 19 free parameters, with more needed for neutrino masses and even more if *supersymmetry* is added.

Result: BSM (Beyond the Standard Model) particle physics has become a major industry.

Grand Unified Theories

Goal: Unify $SU(3)\times SU(2)\times U(1)$ by embedding all three into a single, larger group.

The breaking of the symmetry to $SU(3)\times SU(2)\times U(1)$ is accomplished by introducing new Higgs fields to spontaneously break the symmetry.

In the fundamental theory, before spontaneous symmetry breaking, there is no distinction between an electron, an electron neutrino, or an up or down quark.

The SU(5) Grand Unified Theory

In 1974, Howard Georgi and Sheldon Glashow of Harvard proposed the original grand unified theory, based on $SU(5)$. They pointed out that $SU(3)\times SU(2)\times U(1)$ fits elegantly into $SU(5)$.

To start, let the $SU(3)$ subgroup be matrices of the form

$$u_3 = \begin{pmatrix} x & x & x & 0 & 0 \\ x & x & x & 0 & 0 \\ x & x & x & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} ,$$

where the 3×3 block of x's represents an arbitrary SU(3) matrix.

Similarly let the $SU(2)$ subgroup be matrices of the form

$$u_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \end{pmatrix} ,$$

where this time the 2×2 block of x's represents an arbitrary SU(2) matrix.

Note that u_3 and u_2 commute, since each acts like the identity matrix in the space in which the other is nontrivial.

Finally, the $U(1)$ subgroup can be written as

$$u_1 = \begin{pmatrix} e^{2i\theta} & 0 & 0 & 0 & 0 \\ 0 & e^{2i\theta} & 0 & 0 & 0 \\ 0 & 0 & e^{2i\theta} & 0 & 0 \\ 0 & 0 & 0 & e^{-3i\theta} & 0 \\ 0 & 0 & 0 & 0 & e^{-3i\theta} \end{pmatrix} ,$$

where the factors of 2 and 3 in the exponents were chosen so that the determinant — in this case the product of the diagonal entries — is equal to 1.

Repeating, the $U(1)$ subgroup can be written as

$$u_1 = \begin{pmatrix} e^{2i\theta} & 0 & 0 & 0 & 0 \\ 0 & e^{2i\theta} & 0 & 0 & 0 \\ 0 & 0 & e^{2i\theta} & 0 & 0 \\ 0 & 0 & 0 & e^{-3i\theta} & 0 \\ 0 & 0 & 0 & 0 & e^{-3i\theta} \end{pmatrix} .$$

u_1 commutes with any matrix of the form of u_2 or u_3, since within either the upper 3×3 block or within the lower 2×2 block, u_1 is proportional to the identity matrix.

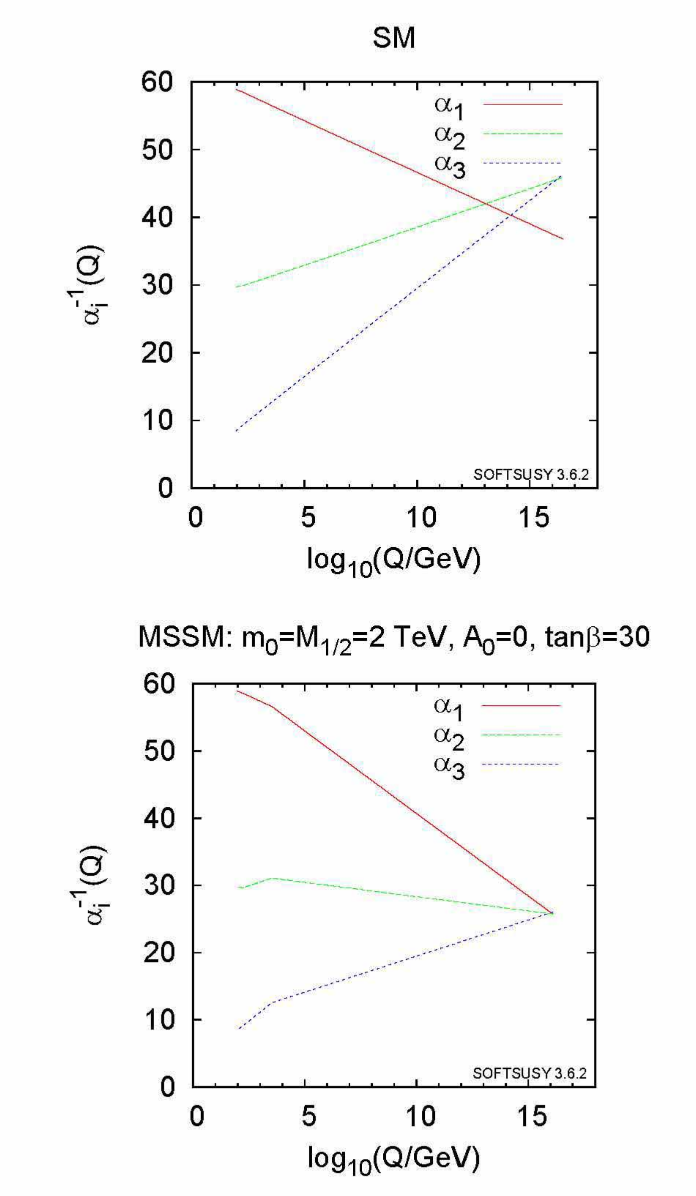
Thus, any element (u_3, u_2, u_1) of $SU(3) \times SU(2) \times U(1)$ can be written as an element u_5 of $SU(5)$, just by setting $u_5 = u_3 u_2 u_1$.

How Can Three Different Types of Interaction Look Like One?

In the standard model, each type of gauge interaction — $SU(3)$, $SU(2)$, and $U(1)$ — has its own interaction strength, described by "coupling constants" g_3, g_2, and g_1. Their values of are different from each other! How can they be one interaction?

BUT: the interaction strength varies with energy in a calculable way. When the calculations are extended to superhigh energies, of the order of 10^{16} GeV, the three interaction strengths become about equal!

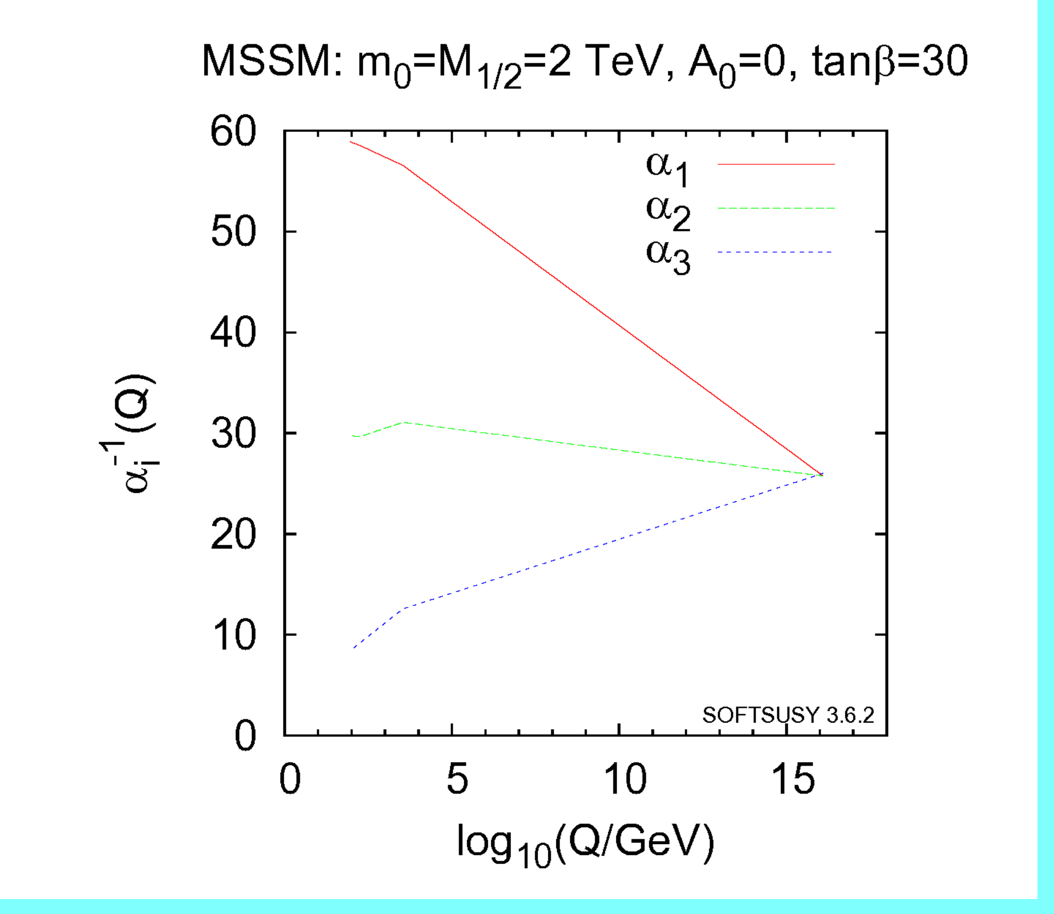
Running Coupling Constants



The top graph shows the running of coupling constants in the standard model, showing that the three coupling constants do not quite meet. The bottom graph shows the running of coupling constants in the MSSM — the Minimal Supersymmetric Standard Model, in which the meeting of the couplings is almost perfect. $\alpha_i = g_i^2/4\pi$.

Source: Particle Data Group 2016 Review of Particle Physics, Chapter 16, *Grand Unified Theories*, Revised January 2016 by A. Hebecker and J. Hisano.

Running Couplings
Minimal Supersymmetric Standard Model



Bottom line: An $SU(5)$ grand unified theory can be constructed by introducing a Higgs field that breaks the $SU(5)$ symmetry to $SU(3)\times SU(2)\times U(1)$ at an energy of about 10^{16} GeV. At energies above 10^{16} GeV, the theory behaves like a fully unified $SU(5)$ gauge theory. At lower energies, it behaves like the standard model. The gauge particles that are part of $SU(5)$ but not part of $SU(3)\times SU(2)\times U(1)$ acquire masses of order 10^{16} GeV.

GUTs (Grand Unified Theories) allow two unique phenomena at low energies, neither of which have been seen:

1) Proton decay. The superheavy gauge particles can mediate proton decay. The minimal $SU(5)$ model — with the simplest conceivable particle content — predicts a proton lifetime of about 10^{31} years, which is ruled out by experiments, which imply a lifetime $\gtrsim 10^{34}$ years.

2) Magnetic monopoles. All grand unified theories imply that magnetic monopoles should be a possible kind of particle. None have been seen.

The absence of evidence does not imply that GUTs are wrong, but we don't know.

8.286 Class 25
December 5, 2022

GRAND UNIFIED THEORIES AND THE MAGNETIC MONOPOLE PROBLEM, PART 3

COSMIC INFLATION!

Announcements

☆ Quiz 3: Wednesday (Walker Memorial, Room 50-340)

- Office hours:

  Today, 5:15 pm – 6:15 pm, Room 8-320, by me

  Tomorrow, 5:00 – 6:00 pm, Room 8-316, by Marianne

- Review Session:

  Today, 6:15 pm, Room 8-320, by Marianne

- For details of coverage, see the Review Problems or the class website.

Grand Unified Theories and the Magnetic Monopole Problem

- Standard Model of Particle Physics: gauge theory with symmetry group $SU(3) \times SU(2) \times U(1)$.

- Gauge symmetry: a symmetry described by $u(x)$, where u is an element of the symmetry group, and $x \equiv (\vec{x}, t)$ is the spacetime coordinate. A gauge transformation changes the fields, but not the physics.

- $SU(3)$ describes the strong interactions, carried by 8 types of gluons. $SU(2) \times U(1)$ describes the weak and electromagnetic interactions, carried by the photon, W^+, W^-, and Z.

- Higgs fields: actually a complex doublet, but we mainly talked about a toy model with a real triplet of Higgs fields, $\vec{\phi}$, transforming under $SU(2)$ like an ordinary 3D vector under ordinary 3D rotations. Invariant under $SU(3)$, multiplied by a phase under $U(1)$.

- Spontaneous symmetry breaking: the minimum energy state is when $|\vec{\phi}| = \phi_v \neq 0$, so it must randomly pick out a direction and break the symmetry down to rotations in 1D, or $U(1)$.

- Masses: the nonzero Higgs field values produce restoring forces for some of the other fields, giving them masses. In particular, the force-carrying gauge bosons associated with broken symmetries acquire a mass, while others remain massless.

- **Grand Unified Theories:** Combine $SU(3)$, $SU(2)$, and $U(1)$ of standard model into one group, the simplest being $SU(5)$. The $SU(5)$ is broken by GUT Higgs fields to $SU(3) \times SU(2) \times U(1)$.

- Predictions of GUTs: proton decay, magnetic monopoles. Magnetic monopoles have not be seen, and $\tau_{\text{proton}} \gtrsim 10^{34}$ years.

The Grand Unified Theory Phase Transition

When $kT \gg 10^{16}$ GeV, the Higgs fields of the GUT undergo large fluctuations, and average to zero. The GUT symmetry is unbroken, and the theory behaves as an SU(5) gauge theory.

As kT falls to about 10^{16} GeV, the matter filling the universe would go through a phase transition, in which some of the components of the GUT Higgs field acquire nonzero values in the thermal equilibrium state, breaking the GUT symmetry. The breaking to $SU(3)\times SU(2)\times U(1)$ might occur in one phase transition, or in a series of phase transitions. We'll assume a single phase transition.

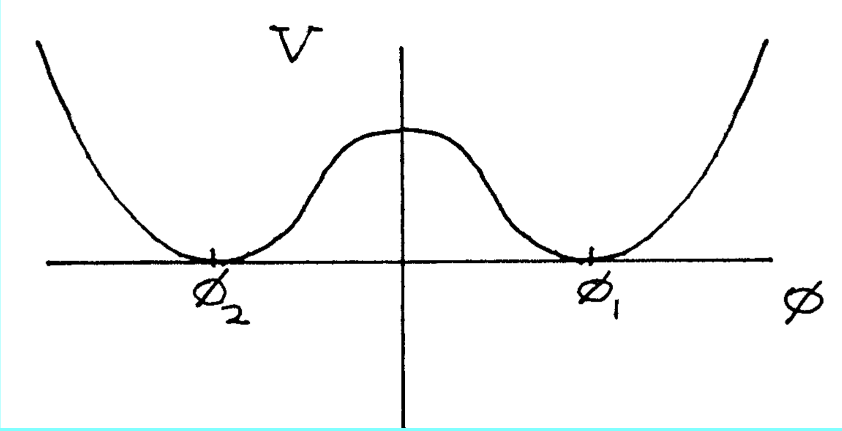
The Higgs fields start to randomly acquire nonzero values, but the nonzero values that form in one region may not align with those in another.

The expression for the energy density contains a term proportional to $|\nabla\Phi|^2$, so the fields tend to fall into low energy states with small gradients. But sometimes the fields in one region acquire a pattern that cannot be smoothly joined with the pattern in a neighboring region, so the smoothing is imperfect, leaving "defects".

Topological Defects

There are three types of defects:

1) Domain walls. For example, imagine a single real scalar field ϕ for which the potential energy function has two local minima, at ϕ_1 and ϕ_2:



Then, as the system cools, some regions will have $\phi \approx \phi_1$ and others will have $\phi \approx \phi_2$. The boundaries between these regions will be surfaces of high energy density: domain walls. Some GUTS allow domain walls, others do not. The energy density of a domain wall is so high that none can exist in the visible universe.

1) Domain walls.

2) Cosmic strings. Linelike defects, which exist in some GUTs but not all.

3) Magnetic monopoles: Pointlike defects, which exist in all GUTs.

Magnetic Monopoles

We'll consider the simplest theory in which monopoles arise, which is exactly the toy vector Higgs model that we have been discussing. It has a 3-component (real) Higgs field, ϕ_a, where $a = 1, 2$ or 3. Gauge symmetry acting on ϕ_a has the same mathematical form as the rotations of an ordinary Cartesian 3-vector.

To be gauge-invariant, the energy density function can depend only on

$$|\phi| \equiv \sqrt{\phi_1^2 + \phi_2^2 + \phi_3^2} \ ,$$

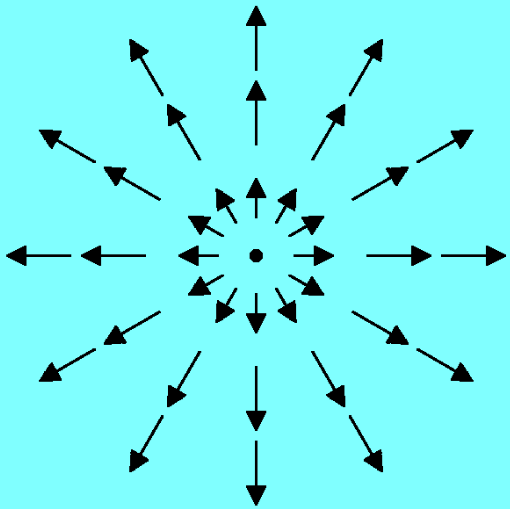
and we assume that it looks qualitatively like the graph for the standard model, with a minimum at H_v.

Now consider the following static configuration,

$$\boxed{\phi_a(\vec{r}\,) = f(r)\hat{r}_a \ ,}$$

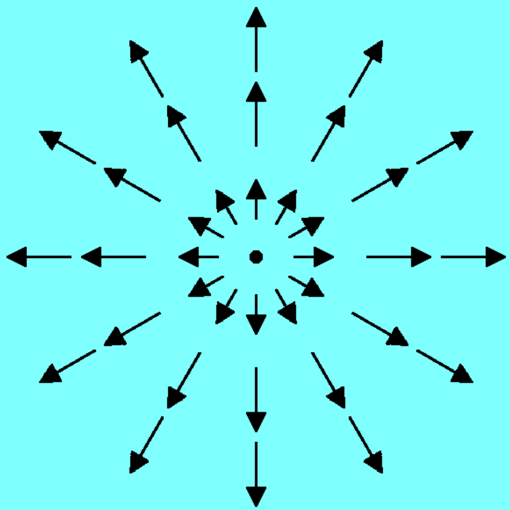
where $r \equiv |\vec{r}\,|$, $\hat{r}_a$ denotes the a-component of the unit vector $\hat{r} = \vec{r}/r$, and $f(r)$ is a function which vanishes when $r = 0$ and approaches H_v as $r \to \infty$.

Pictorially,



where the 3 components of the arrow at each point describe the 3 Higgs field components.

Repeating,



where the 3 components of the arrow at each point describe the 3 Higgs field components.

The directions in gauge space ϕ_a really have nothing to do with directions in physical space, but there is nothing that prevents the fields from existing in this configuration.

The configuration is topologically stable in the following sense: if the boundary conditions at infinity are fixed, and the fields are continuous, then there is always at least one point where $\phi_1 = \phi_2 = \phi_3 = 0$.

Thus, the monopoles are topologically stable knots in the Higgs field.

Why Are These Things Magnetic Monopoles?

Definition: A magnetic monopole is an object with a net magnetic charge, north or south, with a radial magnetic field of the same form as the electric field of a point charge.

Known magnets are always dipoles, with a north end and a south end. If such a magnet is cut in half, one gets two dipoles, each with a north and south end.

Energy of the configuration: the energy density contains a term $\sum_a \vec{\nabla}\phi_a \cdot \vec{\nabla}\phi_a$, but the changing direction of ϕ_a (always radially outward) implies $|\nabla\phi_a| \propto 1/r$. The total energy in a sphere of radius R is proportional to

$$4\pi \int^R r^2 \mathrm{d}r \left(\frac{1}{r}\right)^2 \ ,$$

which diverges as R for large R.

With the vector gauge fields, however, the energy density is more complicated. It can be made finite only if the gauge field configuration corresponds to a net magnetic charge.

Prediction of Magnetic Charge

The magnetic charge is uniquely determined, and is equal to $1/(2\alpha)$ times the electric charge of an electron, where $\alpha \simeq 1/137$ (α = fine-structure constant $= e^2/\hbar c$ in Gaussian units, or $e^2/(4\pi\epsilon_0\hbar c)$ in SI.)

The force between two monopoles is therefore $(68.5)^2$ times as large as the force between two electrons at the same distance. I.e., large!

Kibble Estimate of Magnetic Monopole Production

Since magnetic monopoles are knots in the GUT Higgs fields, they form at the GUT phase transition, when the Higgs fields acquire nonzero mean values. ("Mean" = average over time, not space.)

The density of these knots will be related to the misalignment of the Higgs field in different regions.

Define a correlation length ξ, crudely, as the minimum distance such that the Higgs field at a point is almost uncorrelated with the Higgs field a distance ξ away.

T.W.B. Kibble of Imperial College (London) proposed that the number density of magnetic monopoles (and antimonopoles) can be estimated as

$$n_M \approx 1/\xi^3 \ .$$

Estimate of Correlation Length ξ

In the context of conventional (non-inflationary) cosmology, we can assume

1) that the Higgs field well before the GUT phase transition is in a thermal state, with no long-range correlations.

2) that the universe before the phase transition is well-approximated by a flat radiation-dominated Friedman-Robertson-Walker description.

3) phase transition happens promptly when the temperature of the GUT phase transition is reached, at $kT \approx 10^{16}$ GeV.

Under these assumptions, we are confident that the correlation length ξ must be less than or equal to the horizon length at the time of the phase transition. This seemingly mild limit turns out to have huge implications.

On Problem Set 10, you will calculate the contribution to Ω today, from the monopoles. I won't give away the answer, but you should find that it is greater than 10^{20}.

8.286 Lecture 25 (Part 2)
December 5, 2022

THE INFLATIONARY UNIVERSE

The Inflationary Universe Scenario

- ☆ Inflationary cosmology attempts to describe the behavior of the universe at ridiculously early times — perhaps as early as 10^{-37} seconds.
- ☆ Surprisingly, it can still make predictions that can be tested today.
- ☆ Inflation can provide a solution to the horizon problem, the flatness problem, and the magnetic monopole problem.
- ☆ If correct, inflation can even explain the origin of essentially all the matter in the universe. (One has to start with a bit of matter: a few grams!)

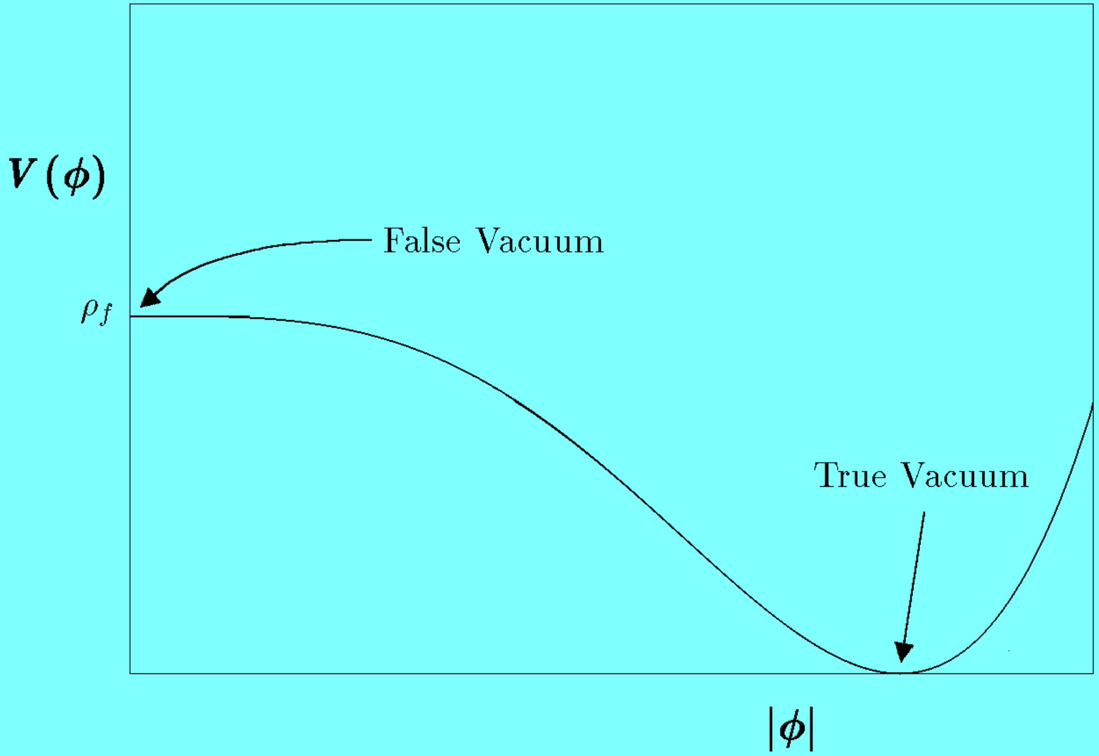
The inflationary scenario assumes the existence of a scalar field ϕ that resembles the Higgs field of the standard model. It is usually assumed to be some beyond-the-standard-model field, associated with a particle of mass $m_\phi c^2$ much higher than anything that we can currently produce in particle accelerators. Any theory with supersymmetry (a symmetry between bosons and fermions), including string theory, leads to many such fields.

Whatever the scalar field that drives inflation is, it is called the “inflaton”.

Inflation is not really a theory, but rather a class of theories, since there are many options for how the inflaton field might behave.

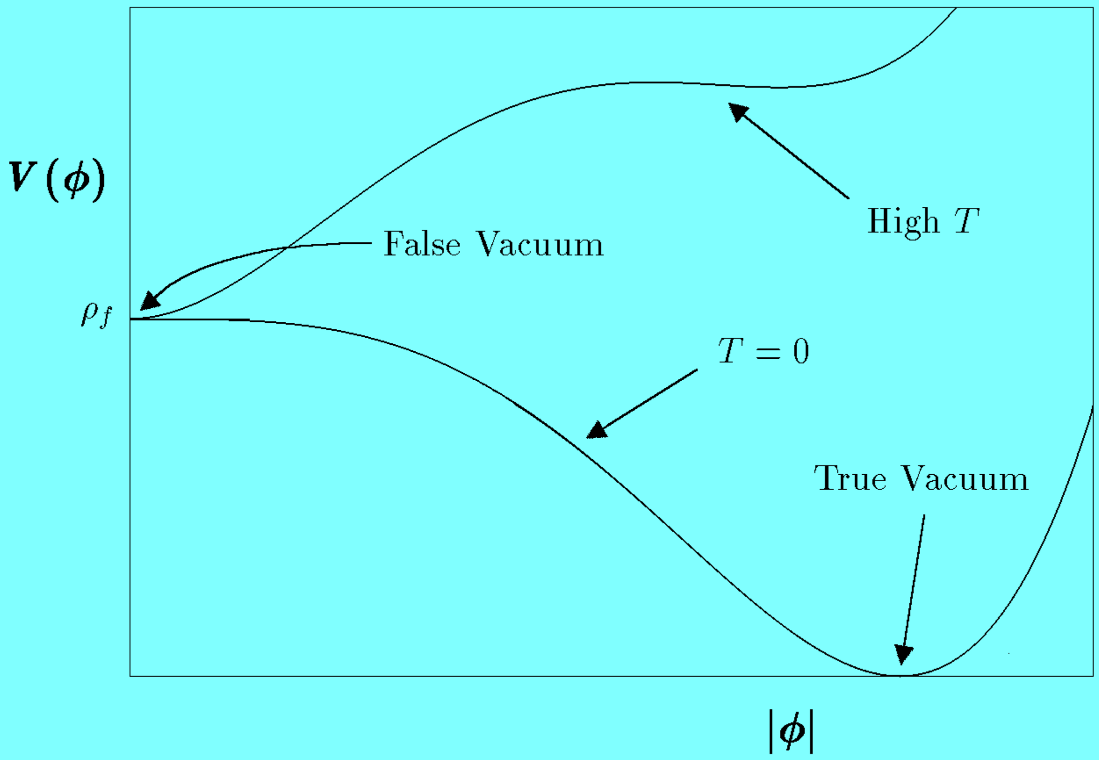
It is conceivable that the inflaton might be the Higgs field of the standard model, but that can work only if the Higgs field interacts with gravity in a particular way, which can be tested only at energies well beyond what we have access to.

The easiest version of inflation to explain is called "hilltop" inflation, or "new" inflation. It assumes an inflaton potential energy density resembling that of the standard model Higgs field:



More general potential energy functions are possible, as we will discuss in a few minutes.

One can also calculate the "finite temperature effective potential" for this theory:



It is the finite temperature effective potential that would be minimized in thermal equilibrium.

Start of Inflation

There is no accepted (or even persuasive) theory of the origin of the universe, so the starting point is uncertain. Inflation starts when the scalar field is at the top of the hill, no matter how it got there.

The scalar field can reach the top of the hill by:

1) Cooling from high temperature ("new" inflation: Linde 1982, Albrecht & Steinhardt, 1982). But: there is not enough time for thermal equilibrium to be reached, so it must be *assumed*.

2) With spatially dependent "chaotic" initial conditions, it will happen somewhere (Linde, 1983). This is probably the dominant point of view today.

3) Creation of the universe by "tunneling from nothing" (Vilenkin, 1983, Linde 1984).

4) Initial conditions for the "wave function of the universe" (Hartle & Hawking, 1983).

5) Who knows?

The good news is that the predictions of inflation do not depend on how it started. This is also bad news, since it means that it is very hard to learn anything about how it started.

The Inflationary Era

Once the inflaton is at the top of the hill, the mass/energy density is fixed, leading to a large negative pressure and gravitational repulsion:

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \; ; \quad \dot{\rho} = 0 \quad \Longrightarrow \quad p = -\rho c^2 \; .$$

Assuming approximate Friedmann-Robertson-Walker evolution,

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right) = \frac{8\pi}{3}G\rho_f ,$$

where ρ_f = mass density of the false vacuum. Thus, ρ_f produces gravitational repulsion.

The de Sitter Solution

The homogeneous isotropic solution can be described as a Robertson-Walker flat universe:

$$\boxed{ds^2 = -c^2 dt^2 + a^2(t) d\vec{x}^2 \ ,}$$

where

$$\boxed{a(t) \propto e^{\chi t} \ , \ \chi = \sqrt{\frac{8\pi}{3} G \rho_f} \ .}$$

This is called de Sitter spacetime.

By a change of coordinates, de Sitter spacetime can, surprisingly, be described as an open universe, a closed universe, or a static universe!

8.286 Class 27
December 12, 2022

COSMIC INFLATION
PART 2

The Inflationary Universe Scenario

- ☆ Inflationary cosmology attempts to describe the behavior of the universe at ridiculously early times — perhaps as early as 10^{-37} seconds.
- ☆ Surprisingly, it can still make predictions that can be tested today.
- ☆ Inflation can provide a solution to the horizon problem, the flatness problem, and the magnetic monopole problem.
- ☆ If correct, inflation can even explain the origin of essentially all the matter in the universe. (One has to start with a bit of matter: a few grams!)

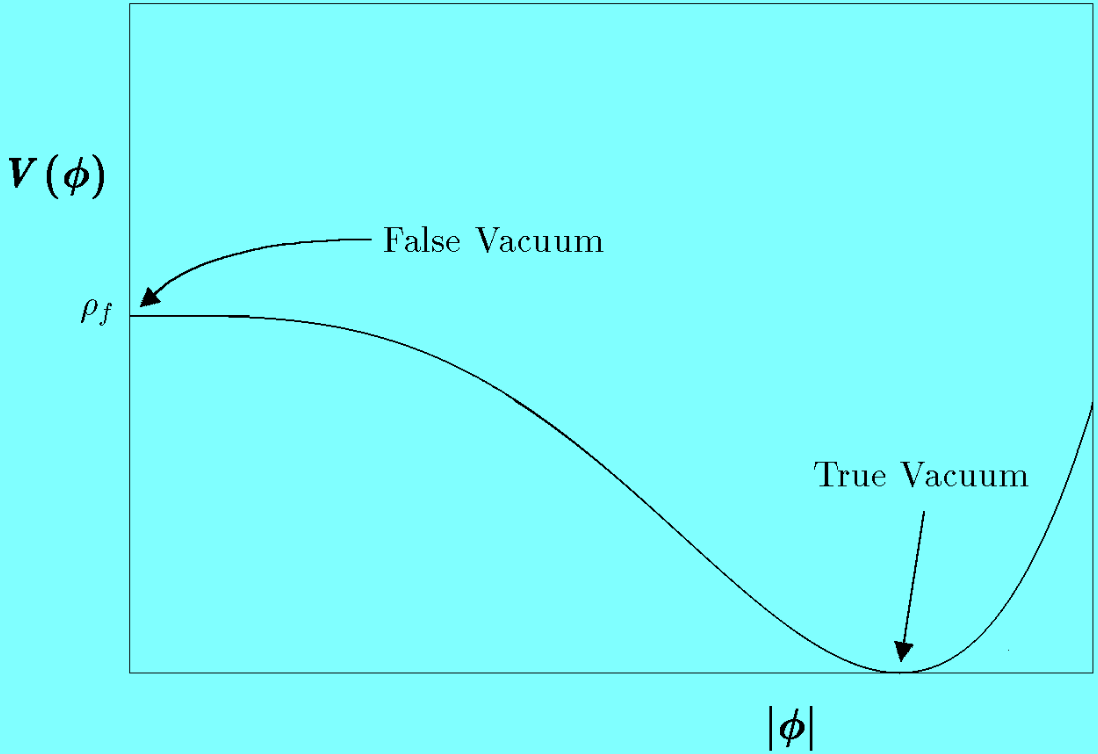
The inflationary scenario assumes the existence of a scalar field ϕ that resembles the Higgs field of the standard model. It is usually assumed to be some beyond-the-standard-model field, very likely at the GUT scale.

Whatever the scalar field that drives inflation is, it is called the "inflaton".

Inflation is not really a theory, but rather a class of theories, since there are many options for how the inflaton field might behave.

It is conceivable that the inflaton might be the Higgs field of the standard model, but that can work only if the Higgs field interacts with gravity in a particular way, which can be tested only at energies well beyond what we have access to.

The easiest version of inflation to explain is called "hilltop" inflation, or "new" inflation. It assumes an inflaton potential energy density resembling that of the standard model Higgs field:



More general potential energy functions are possible, as we will discuss in a few minutes.

Start of Inflation (Summary)

There is no accepted (or even persuasive) theory of the origin of the universe, so the starting point is uncertain. Inflation starts when the scalar field is at the top of the hill, no matter how it got there.

The scalar field can reach the top of the hill by cooling from high temperature, from "chaotic" initial conditions, from "tunneling from nothing," from the Hartle-Hawking "wave function of the universe," or maybe something totally unknown.

Start of Inflation (Summary)

There is no accepted (or even persuasive) theory of the origin of the universe, so the starting point is uncertain. Inflation starts when the scalar field is at the top of the hill, no matter how it got there.

The scalar field can reach the top of the hill by cooling from high temperature, from "chaotic" initial conditions, from "tunneling from nothing," from the Hartle-Hawking "wave function of the universe," or maybe something totally unknown.

The good news is that the predictions of inflation do not depend on how it started.

Start of Inflation (Summary)

There is no accepted (or even persuasive) theory of the origin of the universe, so the starting point is uncertain. Inflation starts when the scalar field is at the top of the hill, no matter how it got there.

The scalar field can reach the top of the hill by cooling from high temperature, from "chaotic" initial conditions, from "tunneling from nothing," from the Hartle-Hawking "wave function of the universe," or maybe something totally unknown.

The good news is that the predictions of inflation do not depend on how it started. This is also bad news, since it means that it is very hard to learn anything about how it started.

The Inflationary Era

Once the inflaton is at the top of the hill, the mass/energy density is fixed, leading to a large negative pressure and gravitational repulsion:

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) \;;\quad \dot{\rho} = 0 \quad \Longrightarrow \quad p = -\rho c^2 \;.$$

Assuming approximate Friedmann-Robertson-Walker evolution,

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right) = \frac{8\pi}{3}G\rho_f ,$$

where ρ_f = mass density of the false vacuum. Thus, ρ_f produces gravitational repulsion.

The de Sitter Solution

The homogeneous isotropic solution can be described as a Robertson-Walker flat universe:

$$\boxed{ds^2 = -c^2 dt^2 + a^2(t) d\vec{x}^2 \ ,}$$

where

$$\boxed{a(t) \propto e^{\chi t} \ , \ \chi = \sqrt{\frac{8\pi}{3} G \rho_f} \ .}$$

This is called de Sitter spacetime.

By a change of coordinates, de Sitter spacetime can, surprisingly, be described as an open universe, a closed universe, or a static universe!

Cosmological "No-Hair" Conjecture

Conjecture: For "reasonable" initial conditions, even if far from homogeneous and isotropic, $\rho = \rho_f$ implies that the region will approach de Sitter space.

Conjectured by Hawking & Moss (1982). Can be proven for linearized perturbations about de Sitter spacetime (Jensen & Stein-Schabes, 1986, 1987). Was shown by Wald (1983) to hold for a class of very large (but spatially homogeneous) perturbations.

Analogous to the Black Hole No-Hair Theorem, which implies that gravitationally collapsing matter approaches a stationary black hole state that depends only on the mass, angular momentum, and charge.

Qualitative behavior: any distortion of the metric is stretched by the expansion to look smooth and flat. Any initial matter distribution is diluted away by the expansion.

De Sitter Event Horizon

In the de Sitter metric, with $a(t) = be^{\chi t}$, the coordinate distance that light can travel between times t_1 and t_2 is

$$\boxed{\Delta r(t_1, t_2) = \int_{t_1}^{t_2} \frac{c}{a(t)}\,\mathrm{d}t = \frac{c}{b}\int_{t_1}^{t_2} e^{-\chi t}\,\mathrm{d}t = \frac{c}{b\,\chi}\left[e^{-\chi t_1} - e^{-\chi t_2}\right] \ ,}$$

which is bounded as $t_2 \to \infty$. If we multiply by $a(t_1)$ and take the limit,

$$\boxed{\lim_{t_2 \to \infty} a(t_1)\,\Delta r(t_1, t_2) = c\chi^{-1} \ ,}$$

which means that if two objects have a physical separation larger than $c\chi^{-1}$, the Hubble length, at any time t_1, light from the first will never reach the second. This is called an event horizon. Event horizons protect an inflating patch from the rest of the universe: once the patch is large compared to $c\chi^{-1}$, nothing from outside can penetrate further than $c\chi^{-1}$.

Event Horizon in the Universe Today

Our universe today is entering a de Sitter phase, in which the dark energy dominates.

In the Review Problems for Quiz 3, Problem 17, the present event horizon was calculated, finding $z = 1.87$.

That means that events that are happening now (i.e., at the same value of the cosmic time), at distances for which the redshift is larger than 1.87, will NEVER be visible to us or our descendents.

The Ending of Inflation

A standard scalar field in a flat FRW universe obeys the equation of motion:

$$\boxed{\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} - \frac{1}{a^2}\nabla_i^2\phi = -\frac{\partial V}{\partial \phi} \ ,}$$

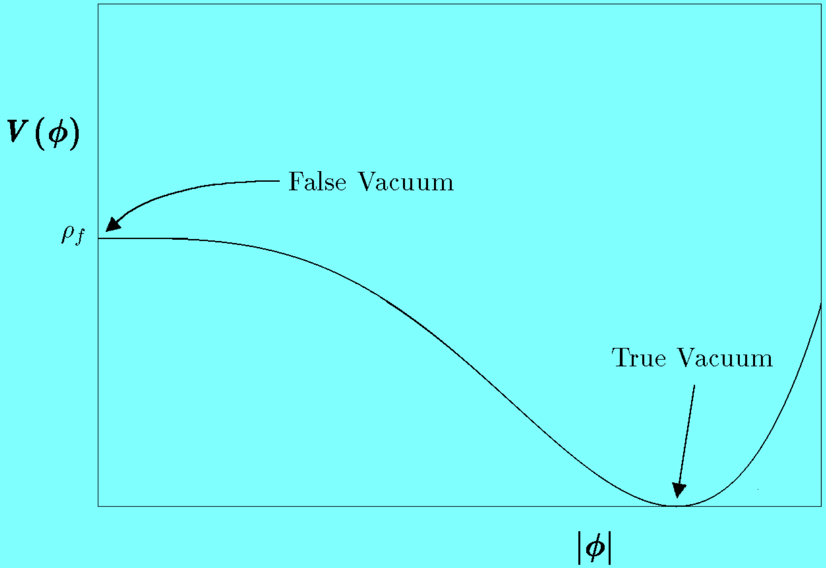
where ∇_i^2 is the Laplacian operator in comoving coordinates x^i, and $V(\phi)$ is the potential energy function (i.e., the potential energy per volume).

The spatial derivative piece soon becomes negligible, due to the $(1/a^2)$ suppression, which reflects the fact that the stretching of space causes ϕ to become nearly uniform over huge regions. The equation is then identical to that of a ball sliding on a hill described by $V(\phi)$, but with a viscous damping (i.e., friction) described by the term $3(\dot{a}/a)\dot{\phi}$.

$$\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} = -\frac{\partial V}{\partial \phi}\ .$$

Fluctuations in ϕ due to thermal and/or quantum effects will cause the field to start to slide down the hill. This will not happen globally, but in regions, typically of size $c\chi^{-1}$.

Within a region, ϕ will start to oscillate about the true vacuum value, at the bottom of the hill. Interactions with other fields will allow ϕ to give its energy to the other fields, producing a "hot soup" of other particles, which is exactly the starting point of the conventional hot big bang theory. This is called *reheating*.



The standard hot big bang scenario begins. Inflation has played the role of a prequel, setting the initial conditions for conventional cosmology.

Numerical Estimates

The energy scale at which inflation happened is not known. One plausible guess is the GUT scale, $E_{\mathrm{GUT}} \approx 10^{16}$ GeV. It cannot be higher (too much gravitational radiation), but can be as low as about 10^3 GeV.

For E_{GUT}, we can estimate

$$\boxed{\rho_f \approx \frac{E_{\mathrm{GUT}}^4}{\hbar^3 c^5} = 2.3 \times 10^{81} \mathrm{g/cm}^3 \ .}$$

Then

$$\boxed{\chi^{-1} \approx 2.8 \times 10^{-38} \text{ s} \ , \quad c\chi^{-1} = 8.3 \times 10^{-28} \text{ cm} \ ,}$$

and the mass of a minimal region of inflation would be about

$$\boxed{M \approx \frac{4\pi}{3} (c\chi^{-1})^3 \rho_f \approx 5.6 \text{ gram.}}$$

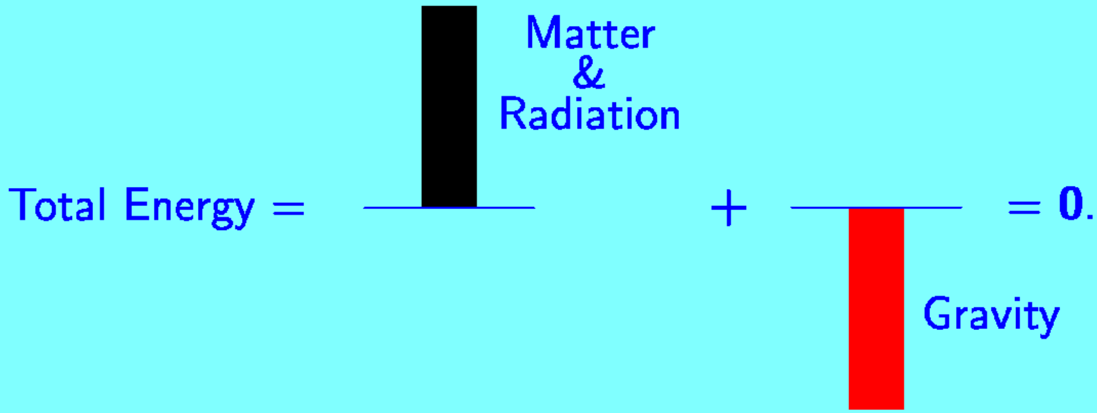
BUT Where Does the Energy Come From?

BUT Where Does the Energy Come From?

- ☆ The energy of a gravitational field is negative (both in Newtonian gravity and in general relativity).
- ☆ The negative energy of gravity cancelled the positive energy of matter, so the total energy was constant and possibly zero.

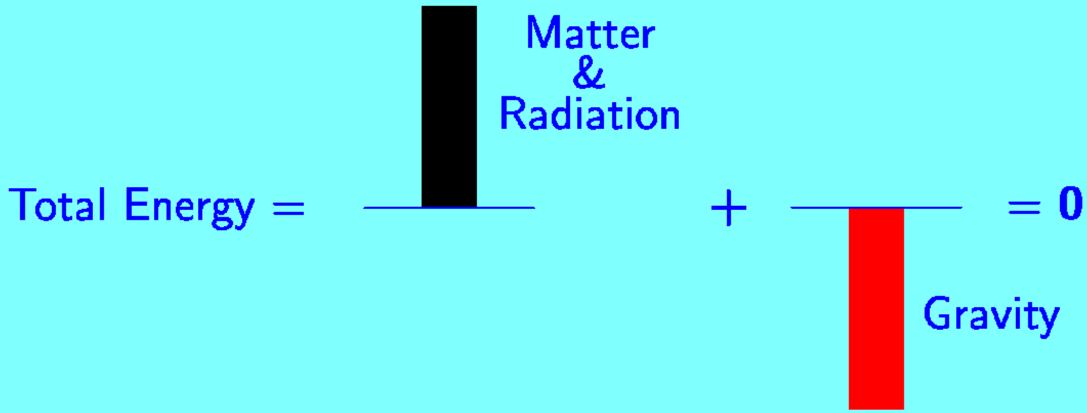
BUT Where Does the Energy Come From?

- ☆ The energy of a gravitational field is negative (both in Newtonian gravity and in general relativity).
- ☆ The negative energy of gravity cancelled the positive energy of matter, so the total energy was constant and possibly zero.
- ☆ The total energy of the universe today is consistent with zero. Schematically,



BUT Where Does the Energy Come From?

- The energy of a gravitational field is negative (both in Newtonian gravity and in general relativity).
- The negative energy of gravity cancelled the positive energy of matter, so the total energy was constant and possibly zero.
- The total energy of the universe today is consistent with zero. Schematically,



- Warning: the concept of total energy in GR is controversial. Some authors would just say that total energy is not defined.

Solutions to the Cosmological Problems

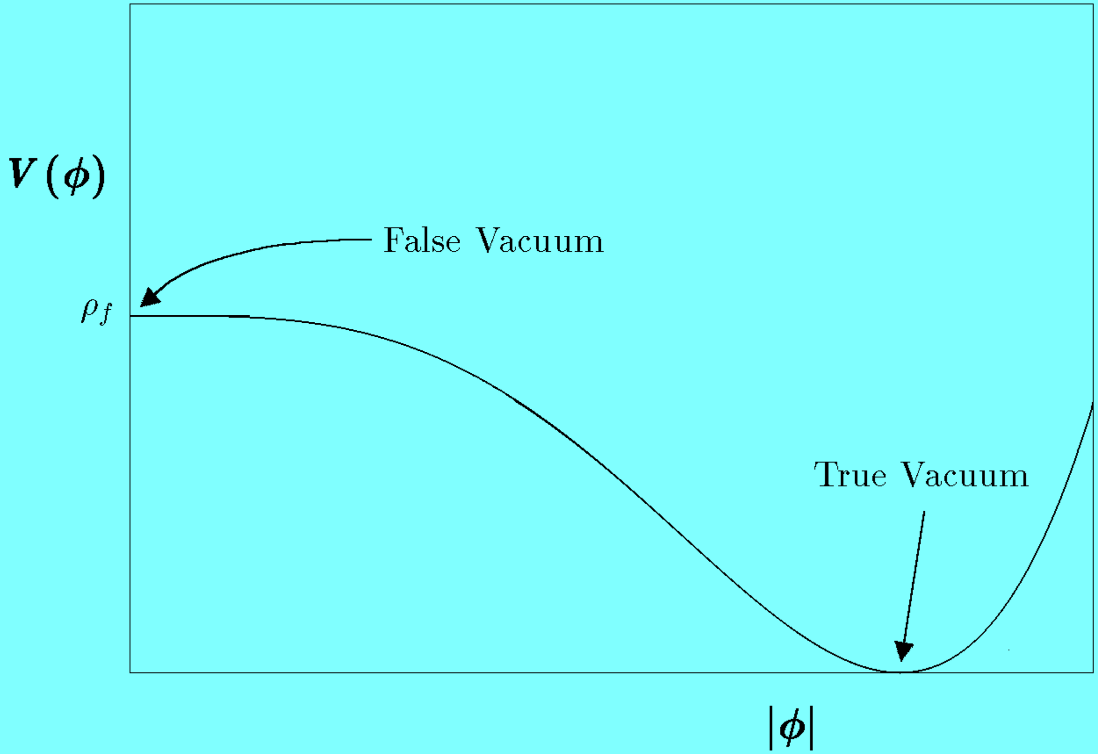
1) **Horizon Problem:** In inflationary models, uniformity is achieved in a tiny region BEFORE inflation starts. Without inflation, such regions would be far too small to matter. But inflation can stretch a tiny region of uniformity to become large enough to include the entire visible universe and more. For inflation at the GUT scale, 10^{16} GeV, we need expansion by about 10^{28}, which is about 65 time constants of the exponential expansion.

8.286 Class 28 (The Last!)
December 14, 2022

COSMIC INFLATION
PART 3

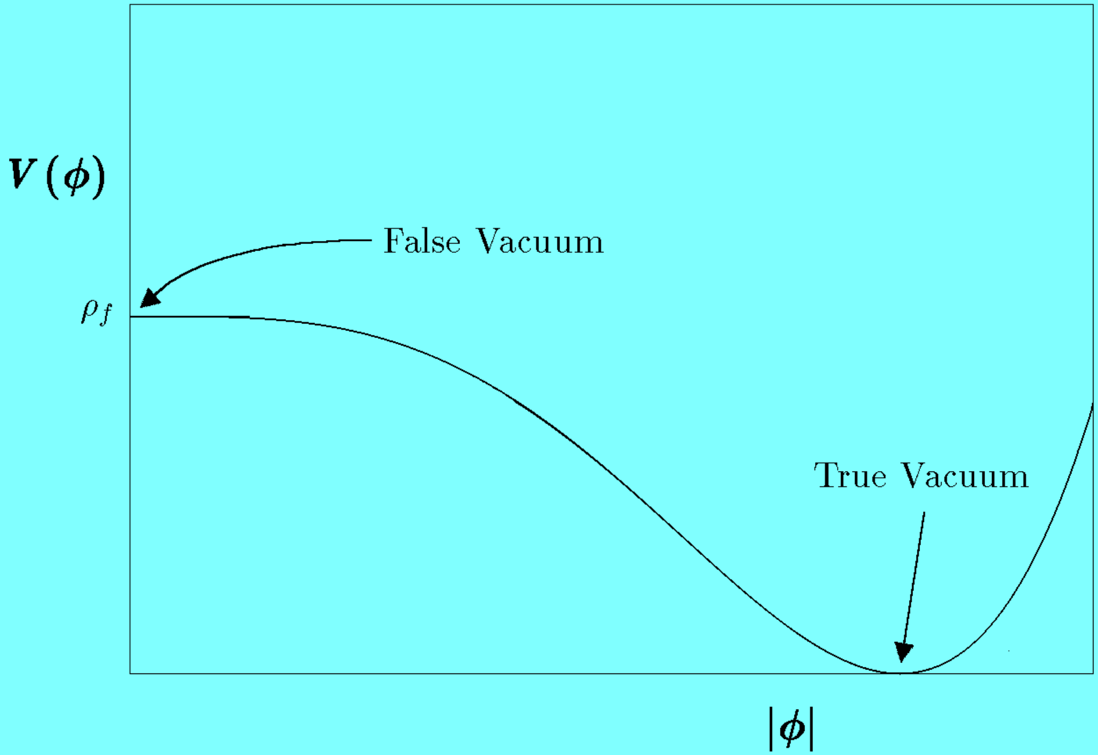
Review from the previous lecture

The easiest version of inflation to explain is called "hilltop" inflation, or "new" inflation. It assumes an inflaton potential energy density resembling that of the standard model Higgs field:



More general potential energy functions are possible, as we will discuss in a few minutes.

The easiest version of inflation to explain is called "hilltop" inflation, or "new" inflation. It assumes an inflaton potential energy density resembling that of the standard model Higgs field:



More general potential energy functions are possible, as we will discuss in a few minutes.

— well, on the next slide.

Not from the last lecture, but it could have been.

Chaotic Inflation

In 1983 Andrei Linde pointed out that inflation can occur in a much more general class of potentials, even one as simple as

$$V(\phi) = \frac{1}{2} m\phi^2 \ .$$

Linde assumed that before inflation the inflaton field ϕ was "chaotic," taking on very different values in different parts of space, so that there will be places where before inflation ϕ was very large.

Linde showed that if the initial value of ϕ was sufficiently large, there would be enough inflation as ϕ rolled down the potential energy hill to accomplish successful inflation.

Summary of Inflationary Scenario

- Initial conditions: We can only speculate about the state of the universe before inflation. But as long as, somehow, the inflaton field in some regions of space was in the range of values that can start inflation, then inflation will happen. Once inflation starts, the inflating region rapidly comes to dominate the volume of the universe.

- Evolution: if ρ is fixed at ρ_f, then $p = -\rho c^2$, and

$$\boxed{\frac{\ddot{a}}{a} = \frac{8\pi}{3} G \rho_f \ .}$$

☆ Evolution: if ρ is fixed at ρ_f, then $p = -\rho c^2$, and

$$\boxed{\frac{\ddot{a}}{a} = \frac{8\pi}{3} G \rho_f \ .}$$

☆ A simple, homogeneous and isotropic solution is the de Sitter solution:

$$\boxed{\mathrm{d}s^2 = -c^2 \mathrm{d}t^2 + a^2(t) \mathrm{d}\vec{x}^2 \ ,}$$

where

$$\boxed{a(t) \propto e^{\chi t} \ , \ \chi = \sqrt{\frac{8\pi}{3} G \rho_{\mathrm{f}}} \ .}$$

☆ A simple, homogeneous and isotropic solution is the de Sitter solution:

$$\boxed{\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\mathrm{d}\vec{x}^2 \ ,}$$

where

$$\boxed{a(t) \propto e^{\chi t} \ , \ \chi = \sqrt{\frac{8\pi}{3}G\rho_{\mathrm{f}}} \ .}$$

☆ Cosmological "No-Hair" Conjecture: For "reasonable" initial conditions, even if far from homogeneous and isotropic, $\rho = \rho_f$ implies that the region will approach de Sitter space.

☆ De Sitter Event Horizon: In de Sitter space, if two objects have a physical separation larger than $c\chi^{-1}$, the Hubble length, at any time t_1, light from the first will never reach the second. This is called an *event horizon*.

The Ending of Inflation

A standard scalar field in a flat FRW universe obeys the equation of motion:

$$\boxed{\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} - \frac{1}{a^2}\nabla_i^2\phi = -\frac{\partial V}{\partial \phi}\ ,}$$

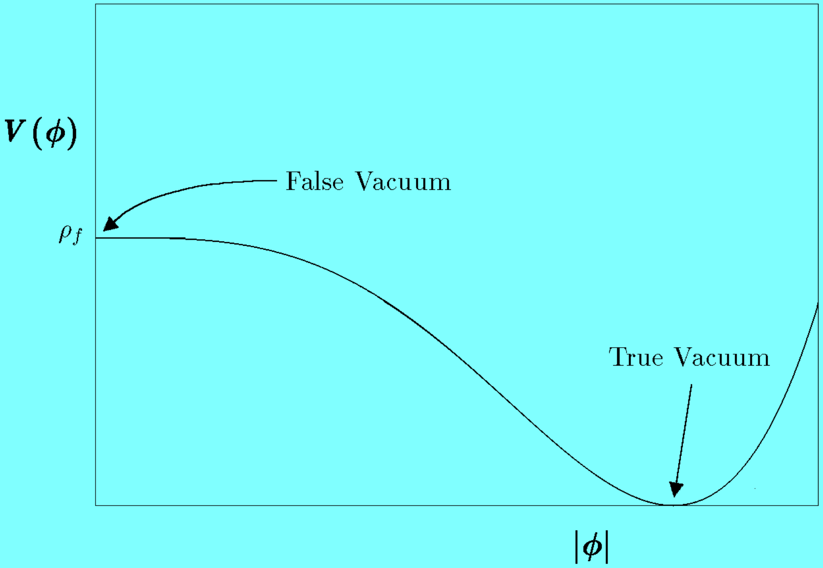
where ∇_i^2 is the Laplacian operator in comoving coordinates x^i, and $V(\phi)$ is the potential energy function (i.e., the potential energy per volume).

The spatial derivative piece soon becomes negligible, due to the $(1/a^2)$ suppression, which reflects the fact that the stretching of space causes ϕ to become nearly uniform over huge regions. The equation is then identical to that of a ball sliding on a hill described by $V(\phi)$, but with a viscous damping (i.e., friction) described by the term $3(\dot{a}/a)\dot{\phi}$.

$$\ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} = -\frac{\partial V}{\partial \phi} \ .$$

Fluctuations in ϕ due to thermal and/or quantum effects will cause the field to start to slide down the hill. This will not happen globally, but in regions, typically of size $c\chi^{-1}$.

Within a region, ϕ will start to oscillate about the true vacuum value, at the bottom of the hill. Interactions with other fields will allow ϕ to give its energy to the other fields, producing a "hot soup" of other particles, which is exactly the starting point of the conventional hot big bang theory. This is called *reheating*.



The standard hot big bang scenario begins. Inflation has played the role of a prequel, setting the initial conditions for conventional cosmology.

Solutions to the Cosmological Problems

1) **Horizon Problem:** In inflationary models, uniformity is achieved in a tiny region BEFORE inflation starts. Without inflation, such regions would be far too small to matter. But inflation can stretch a tiny region of uniformity to become large enough to include the entire visible universe and more. For inflation at the GUT scale, 10^{16} GeV, we need expansion by about 10^{28}, which is about 65 time constants of the exponential expansion.

2) **Flatness Problem:** Just look at Friedmann equation:

$$\boxed{\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2}\ .}$$

"Flatness" is the statement that the final term in this equation is negligible. But during inflation, $\rho \approx \rho_v = \text{const}$, while $a(t)$ grows exponentially. If $a(t)$ grows by at least 10^{28} during inflation, the final term is suppressed by a factor of $\left(10^{28}\right)^2 = 10^{56}$.

3) **Monopole Problem:** Solved by dilution, as long as the inflation occurs during or after the process of monopole production. For inflation at the GUT scale, the volume of any comoving region increases during inflation by a factor of about $\left(10^{28}\right)^3 = 10^{84}$ or more! That is plenty enough to make monopoles impossible to find.

Some small number of monopoles could be produced during reheating, so it makes sense to look for them. But, except for an irreproducible single event seen by Blas Cabrera at Stanford in 1982, magnetic monopoles have not been seen.

Ripples in the Cosmic Microwave Background

The CMB is uniform in all directions to an accuracy of a few parts in 100,000. Nonetheless, at the level of a few parts in 100,000 there ARE anisotropies, and they have now been measured to high precision. Since the CMB is essentially a snapshot of the universe at $t \approx 380,000$ yr, these ripples are interpreted as perturbations in the cosmic mass density at this time.

In the early days of inflation, such density perturbations were a cause for worry. (The ripples had not yet been seen, but cosmologists knew that the early universe must have had density perturbations, or else galaxies and stars could never have formed.) Inflation smooths out the universe so effectively, that it looked like no density perturbations could survive.

Quantum Mechanics to the Rescue (Again)

Why again? We spoke earlier about how quantum mechanics was necessary to save us from freezing to death. If classical mechanics ruled, all thermal energy would gradually disappear into shorter and shorter wavelength electromagnetic radiation.

If inflation happened with classical physics, it would smooth the universe so perfectly that stars and galaxies could never form.

But quantum mechanics is intrinsically probabilistic. While the classical version of inflation predicts an almost exactly uniform mass density, the intrinsic randomness of the quantum version implies that the mass density will be a little higher in some places, and a little lower in others.

In 1965, Andrei Sakharov, the Russian nuclear physicist and political activist, proposed in a rather wildly speculative paper that quantum fluctuations might account for the structure of the universe.

In 1981, Mukhanov and Chibisov tried to calculate the density fluctuations in pre-inflationary/inflationary model invented by Alexei Starobinsky in 1980.

In summer 1982, Gary Gibbons and Stephen Hawking organized the Nuffield Workshop on the Very Early Universe in Cambridge UK, where a number of physicists worked feverishly and argued through the night about how to calculate these perturbations in inflation. In the end, all agreed. Four papers emerged: Hawking, Starobinsky, Guth & Pi, and Bardeen, Steinhardt, & Turner.

Basic conclusion: the amplitude of the density perturbations is very "model-dependent," meaning that it depends on the unknown details of $V(\phi)$. But: the spectrum — the way in which the intensity of the ripples depends on the wavelength of the ripples — is the same for a wide range of "simple" inflationary models. Simple = "Single field / slow-roll models," i.e. models with a single inflaton field, and with small values for $\mathrm{d}V/\mathrm{d}\phi$ and $\mathrm{d}^2V/\mathrm{d}\phi^2$.

Observations of the Ripples in the CMB

In 1982, it seemed (at least to me) out of the question that these ripples would ever be seen.

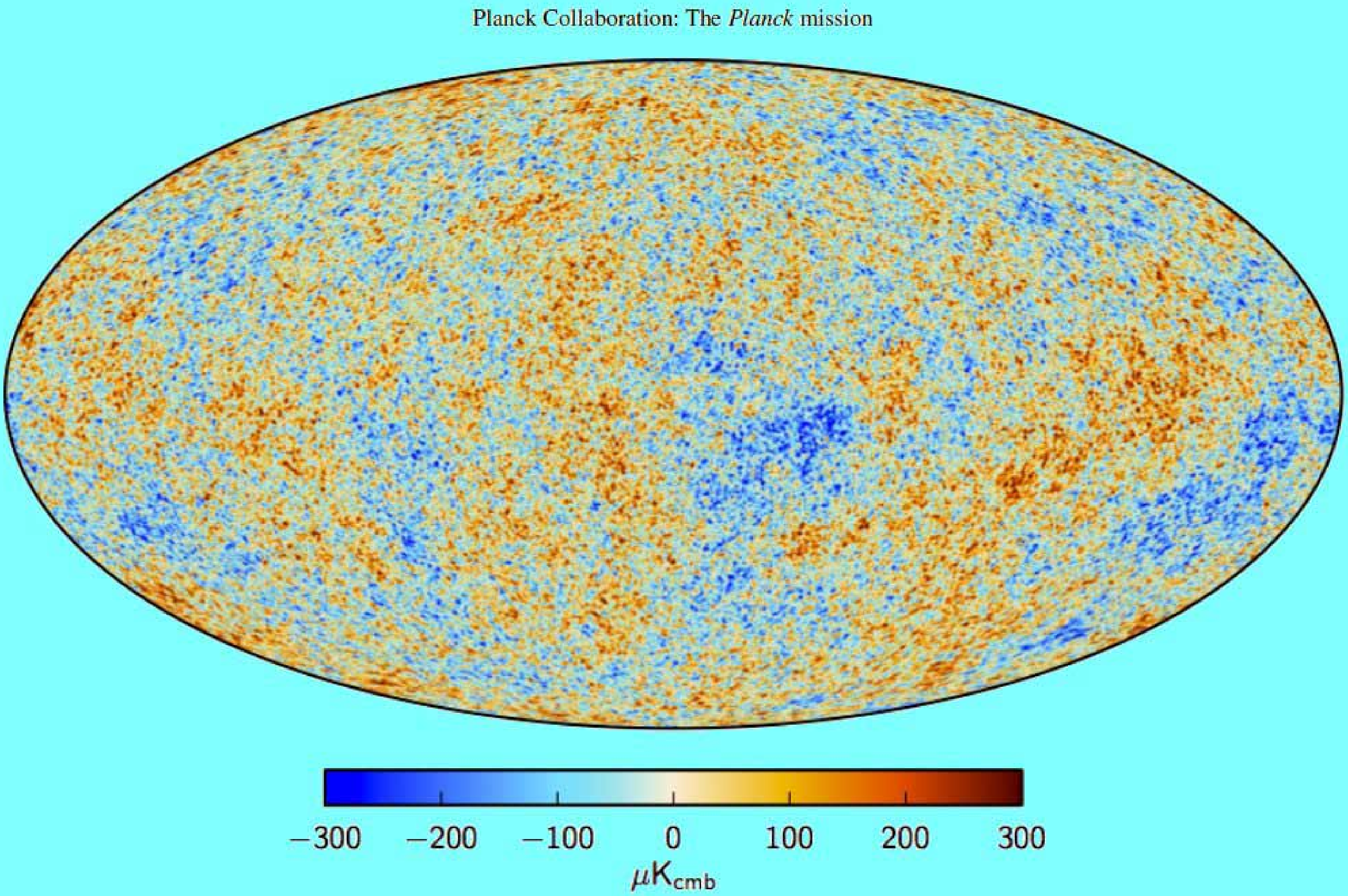
There have now been 3 satellite experiments to measure the CMB, plus many many ground-based experiments. The three satellites were:

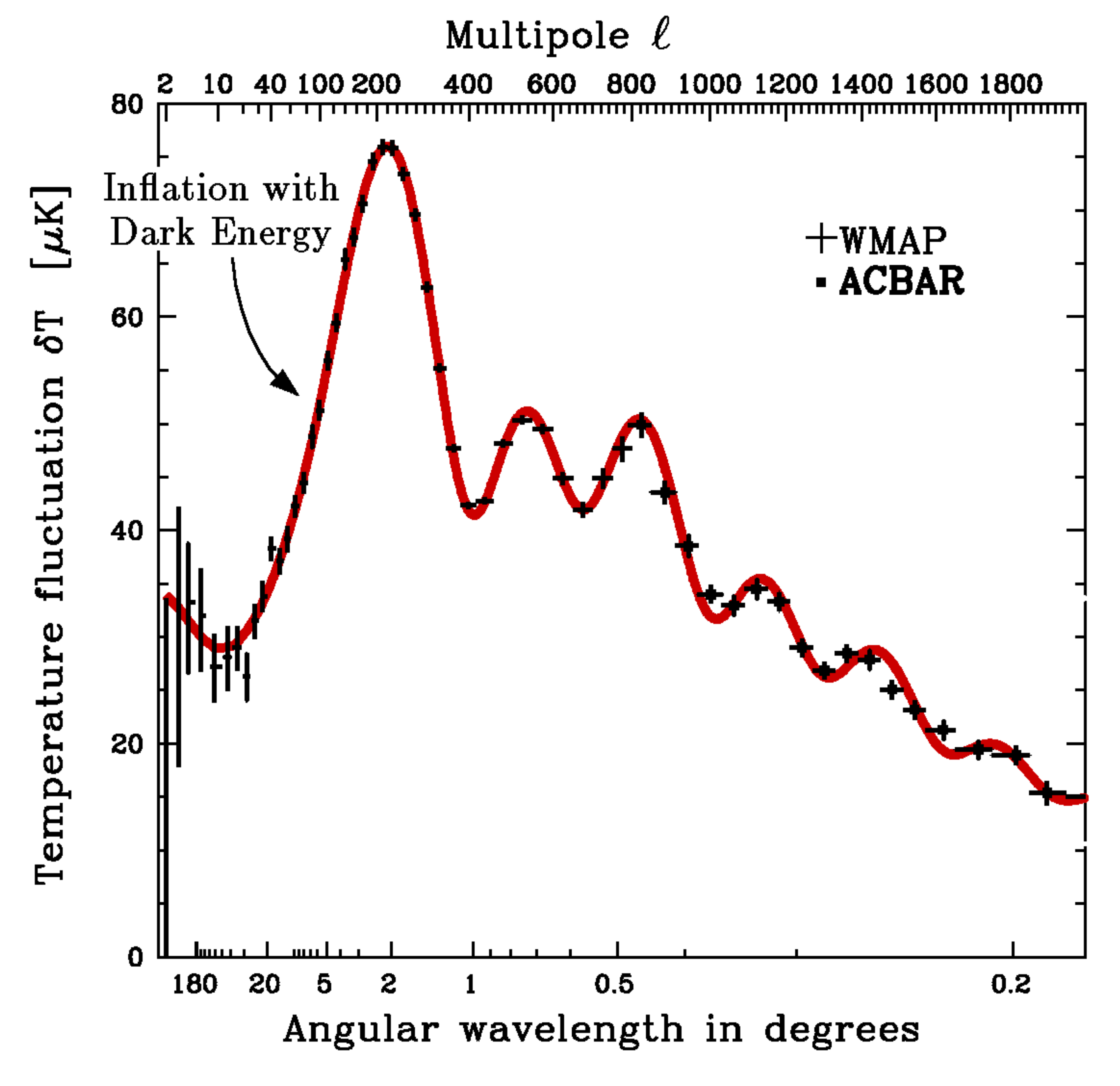
COBE: Cosmic Background Explorer, launched by NASA in 1989, after 15 years of planning. In 1992 it announced its first measurements of CMB anisotropies. The angular resolution was crude, about 7°, but the results agreed with inflation.

WMAP: The Wilkinson Microwave Anisotropy Probe, launched by NASA in 2001. 45 times more sensitive, with 33 times better angular resolution than COBE. Still consistent with inflation.

Planck: Launched in 2009 by ESA. Resolution about 2.5 times better than WMAP. Results still consistent with inflation.

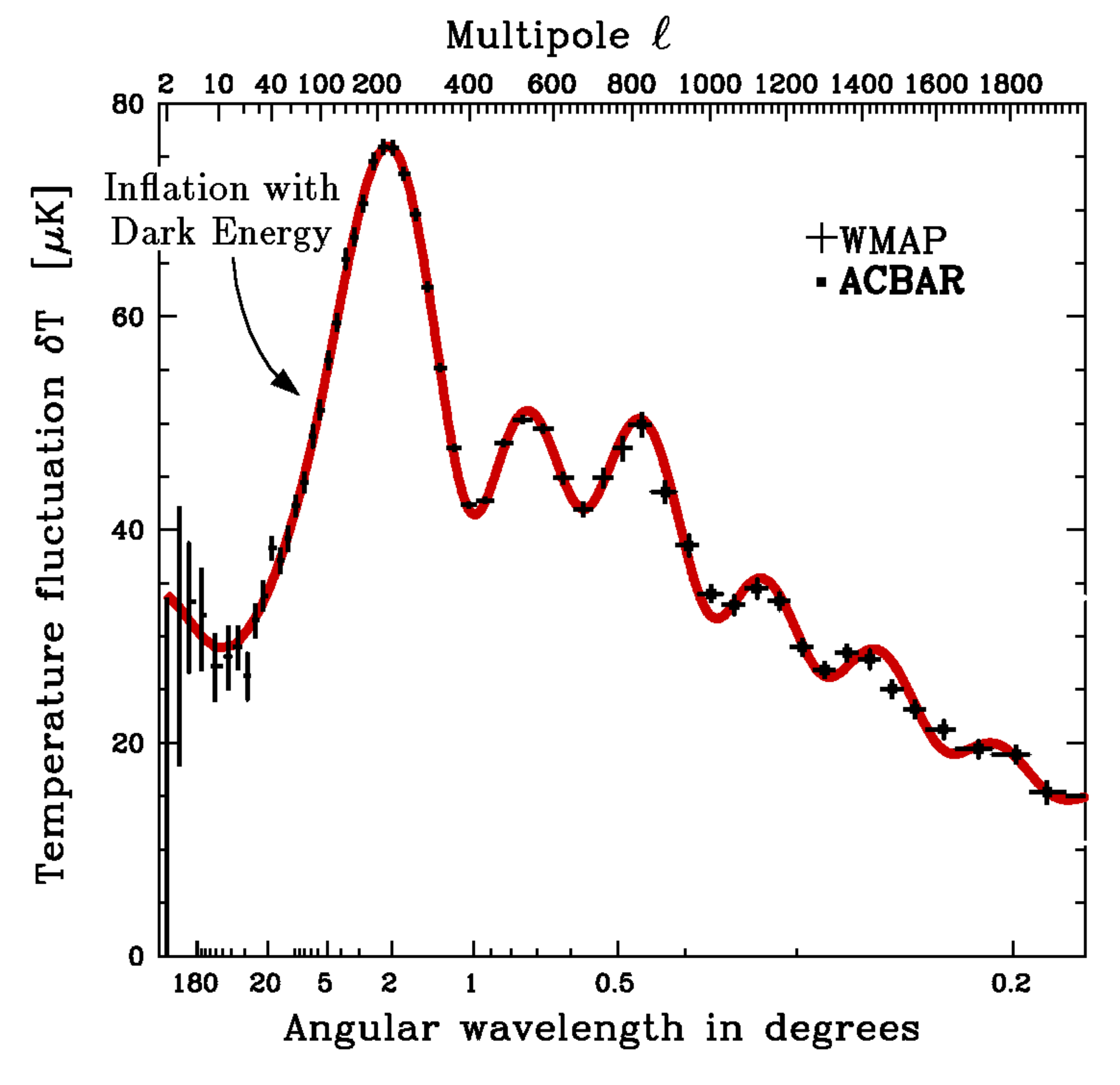
Ripples in the Cosmic Microwave Background





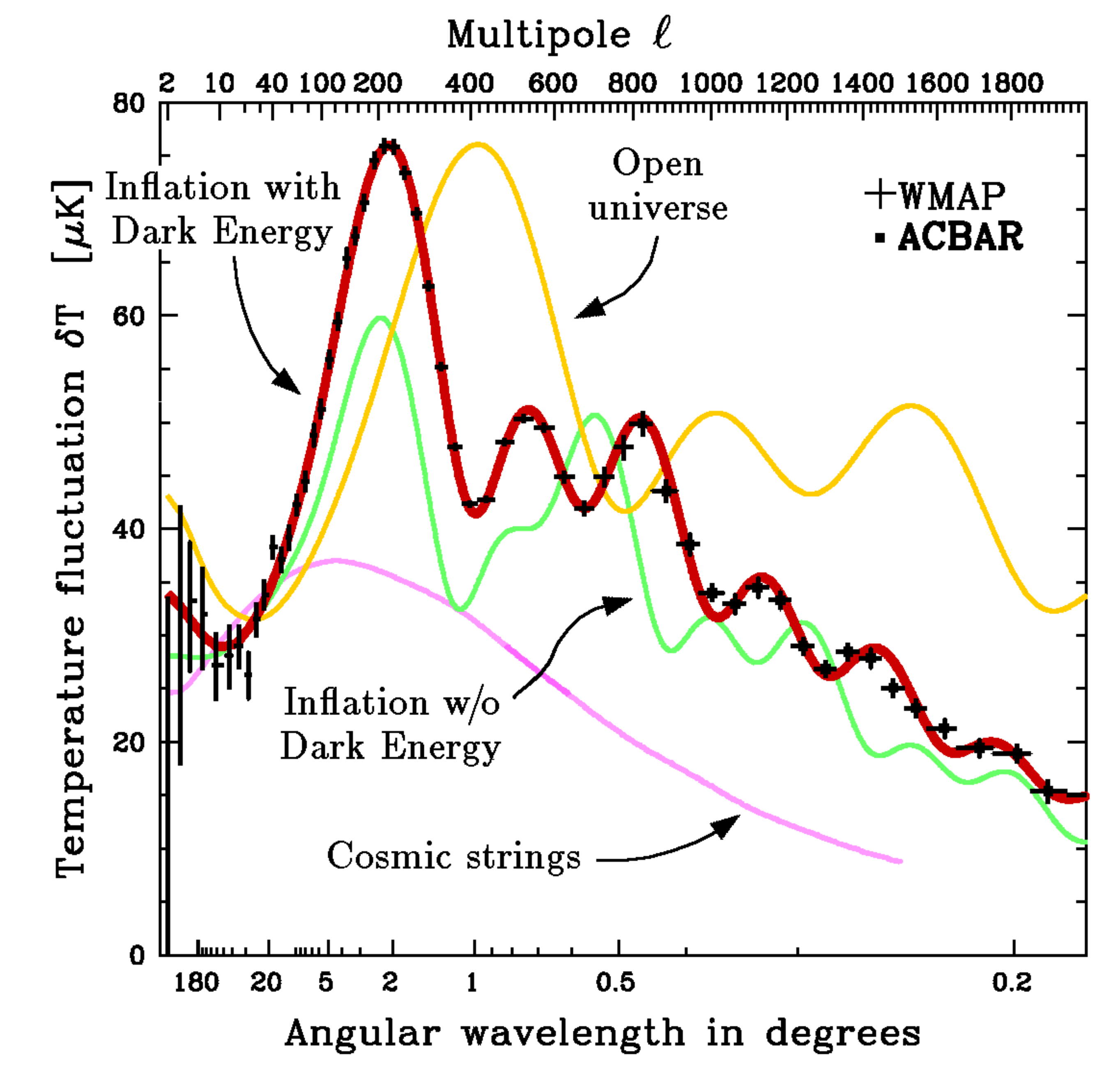
CMB: Comparison of Theory and Experiment

Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data (Jan 2010).



CMB: Comparison of Theory and Experiment

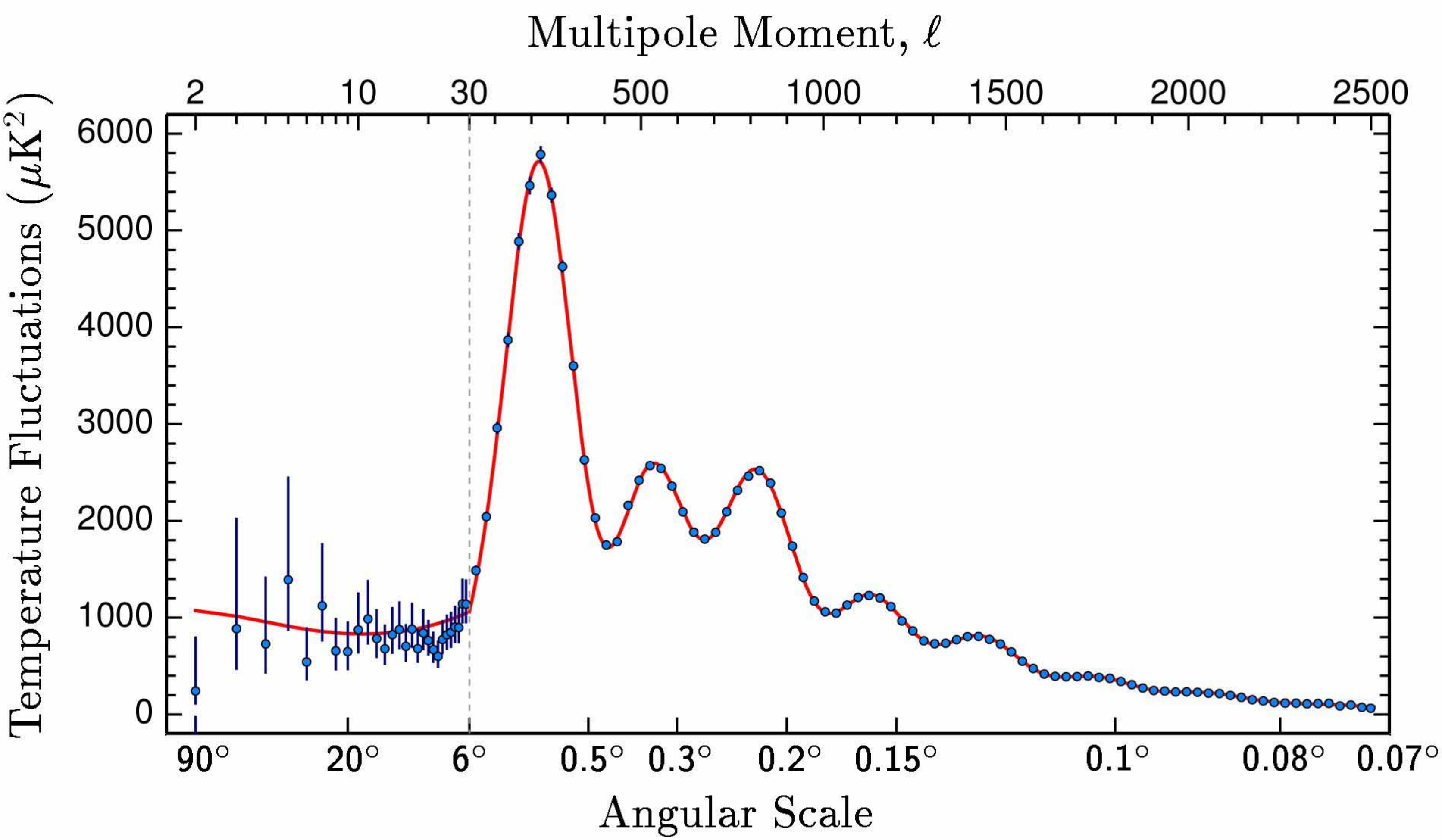
Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data (Jan 2010).



CMB: Comparison of Theory and Experiment

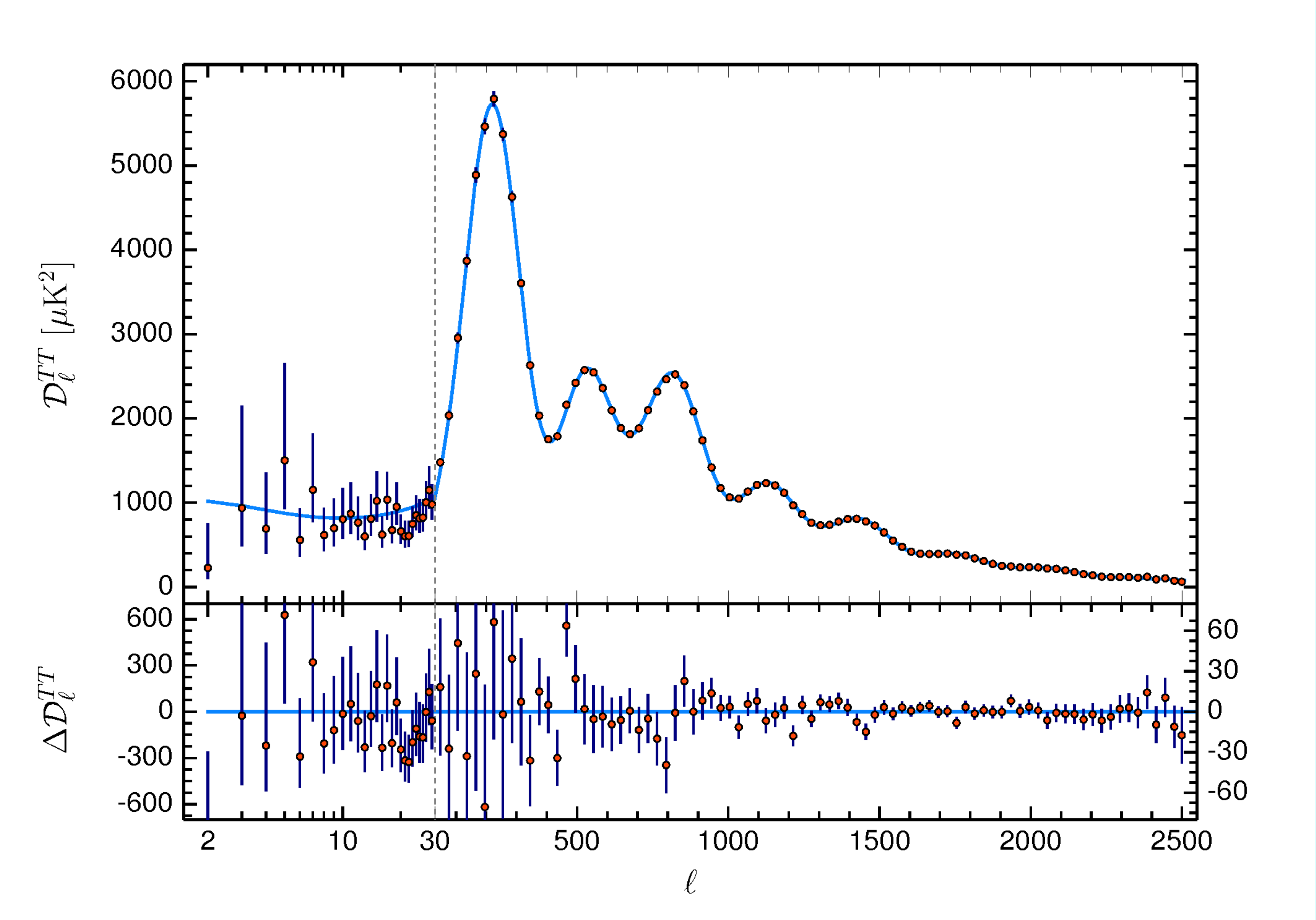
Graph by Max Tegmark, for A. Guth & D. Kaiser, *Science* **307, 884** (Feb 11, 2005), updated to include WMAP 7-year data.

Planck 2015 Spectrum



Planck Collaboration, 2015

Planck 2018 Spectrum



Lower panel shows difference between data and model.

This and the following slides were not reached in class.

Eternal Inflation

In hilltop inflation, while the scalar field rolls down the hill in the potential energy diagram, there is always some small quantum mechanical probability that the field remains at the top.

Approximate calculations show that the probability of remaining at the top falls off exponentially with time. That is, the false vacuum has an exponential decay law, like a radioactive substance.

In any successful model of inflation, the half-life of the false vacuum is much longer than the doubling time of the exponential expansion of $a(t)$.

So, in one half-life of the decay, half of the region in false vacuum stops inflating, but the region remaining in the false vacuum state becomes much larger than the original size of the full region! Thus, the volume of false vacuum region grows exponentially in time.

The ending of inflation happens in localized patches, where in each patch there is a local big bang, forming what we call a "pocket universe". The theory seems to lead to the production of pocket universes ad infinitum. The collection of pocket universes is called a "multiverse".

Is this relevant to physics?

Maybe. It offers a possible explanation of the very small vacuum energy density of our universe. If there is an infinite set of pocket universes, with each one filled with a different vacuum-like state (string theory, for example, gives a huge number of vacuum-like states), then there will be pocket universes with very small vacuum energies. Only those with small vacuum energies will develop life, since the others will implode of fly apart before life could form. All of this is speculative and controversial, however.

Bubble Nucleation in an Eternally Inflating Universe