From theories to observations (how to make computers do physics)

Deanna C. Hooper

(they/them)

Kosmologian kesäkoulu May 2024

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

HELSINKI INSTITUTE OF PHYSICS

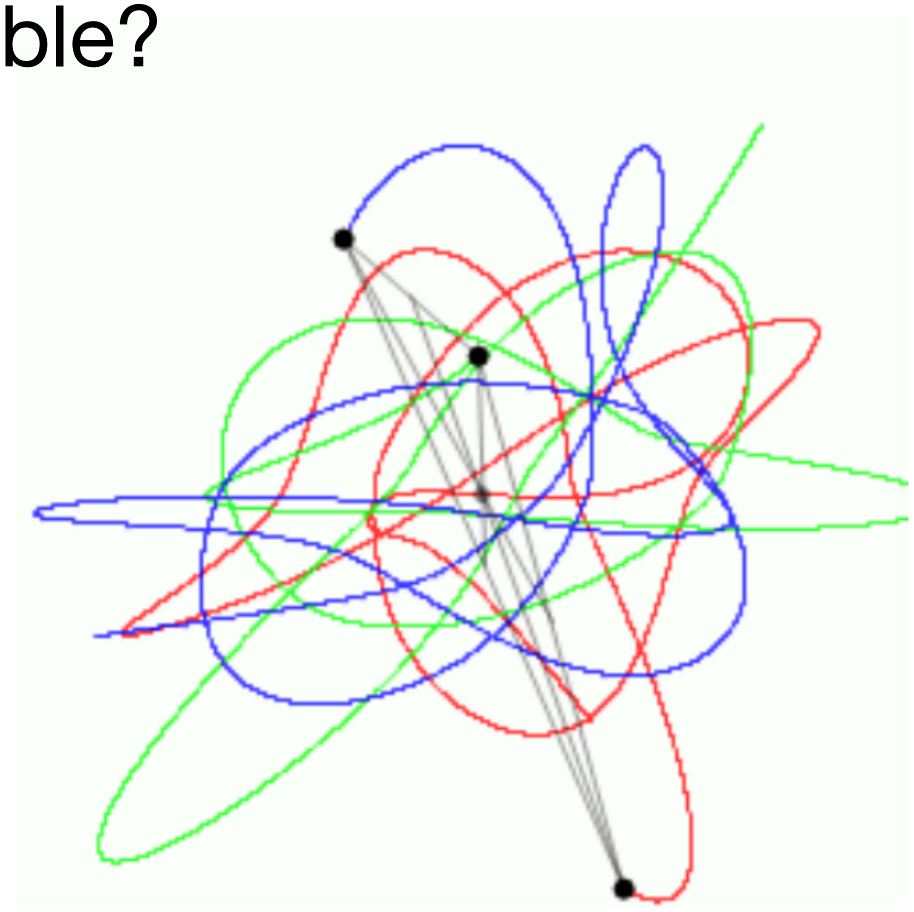
What to expect from this: how to get the computer to do work for you

1. Why we need computational cosmology
2. What type of equations go into our theories
3. How to put equations on a computer
4. Model comparison and statistics

Why we need computational cosmology

There are several problems in cosmology we need to deal with

- We are part of the system we are observing — how do we study the universe from inside the universe?
- We only have one universe in which we can do our experiments — is cosmology reproducible?
- Sometimes the theories rely on very difficult mathematics — for example, the three body problem

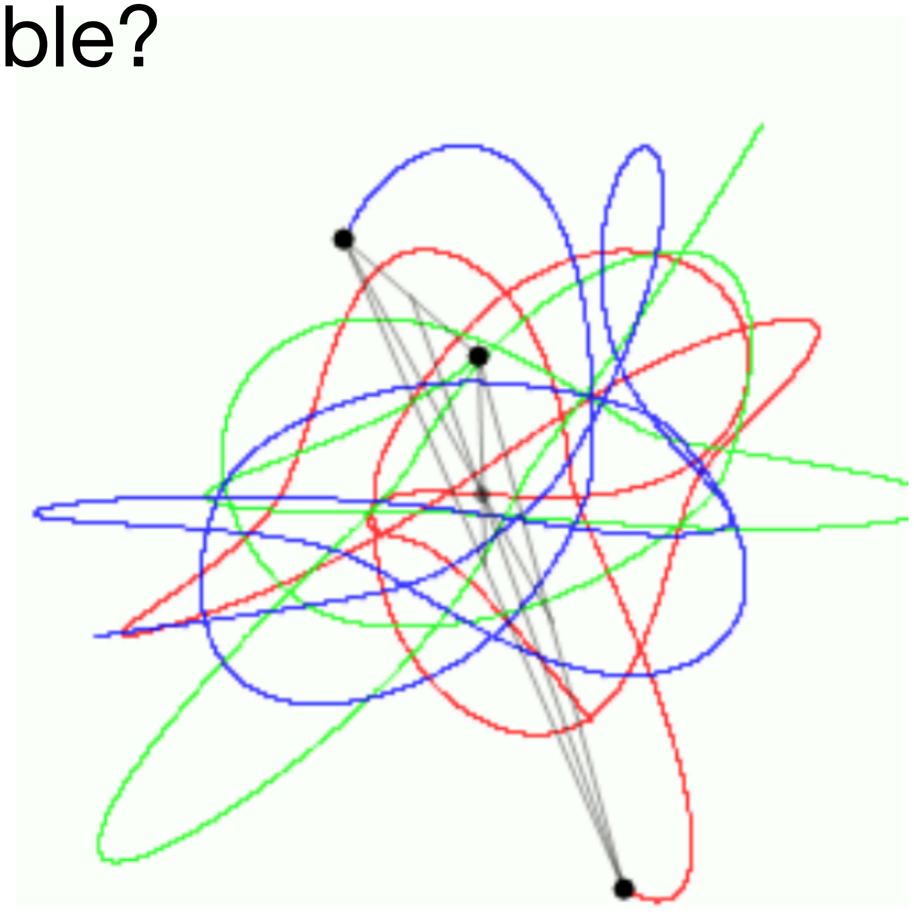


Animation available [here]()

There are several problems in cosmology we need to deal with

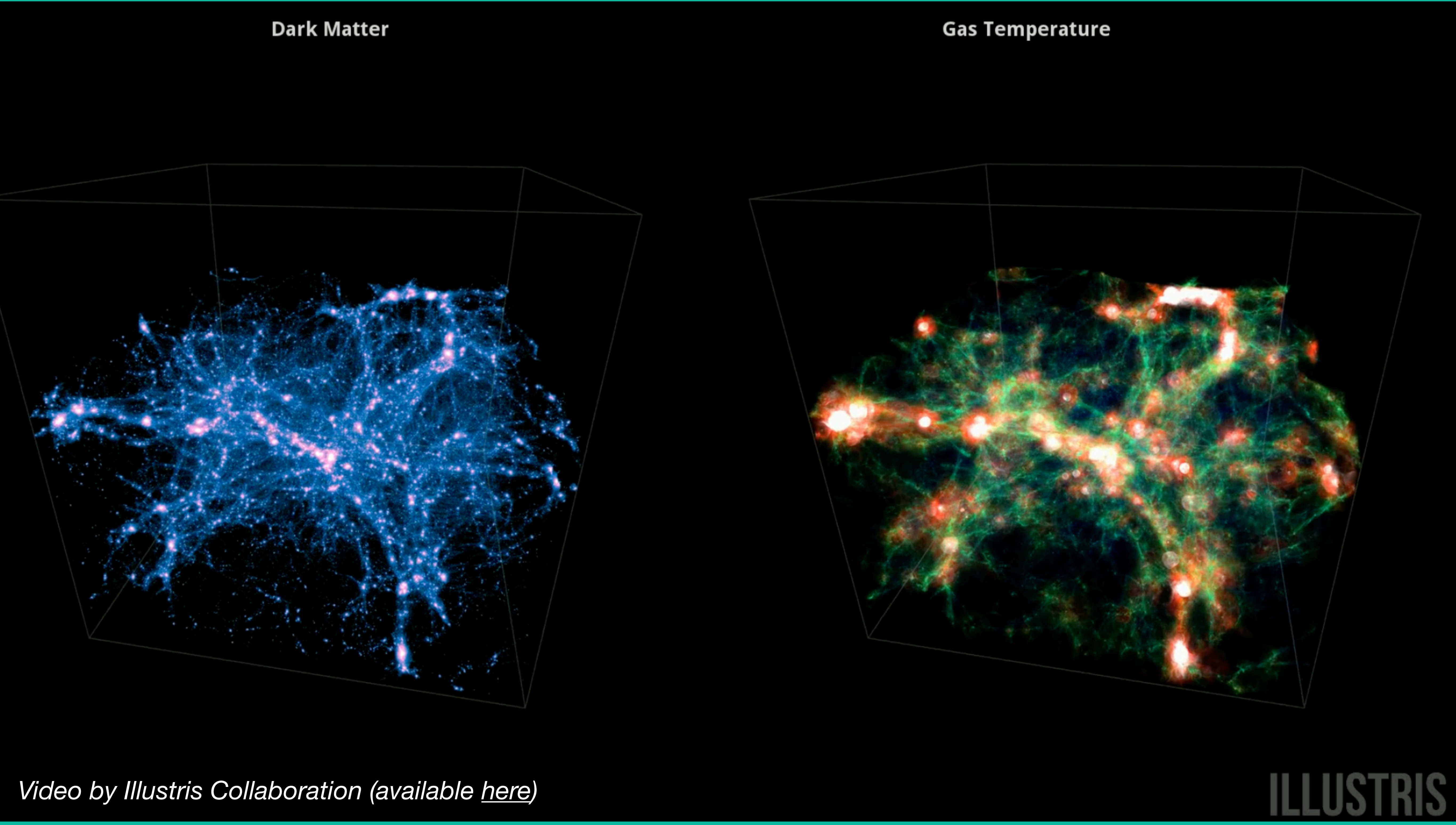
- We are part of the system we are observing — how do we study the universe from inside the universe?
- We only have one universe in which we can do our experiments — is cosmology reproducible?
- Sometimes the theories rely on very difficult mathematics — for example, the three body problem

Would be nice if we could put the universe in a box to do experiments and tests…



Animation available [here]

Come to the dark side, we have pretty pictures (and sometimes cookies)!



Video by Illustris Collaboration (available here)

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

4. Build an experiment to get the data

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

4. Build an experiment to get the data

5. Produce many versions of what the data would look like for different underlying models and parameters

From theories to observations: a step by step guide

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

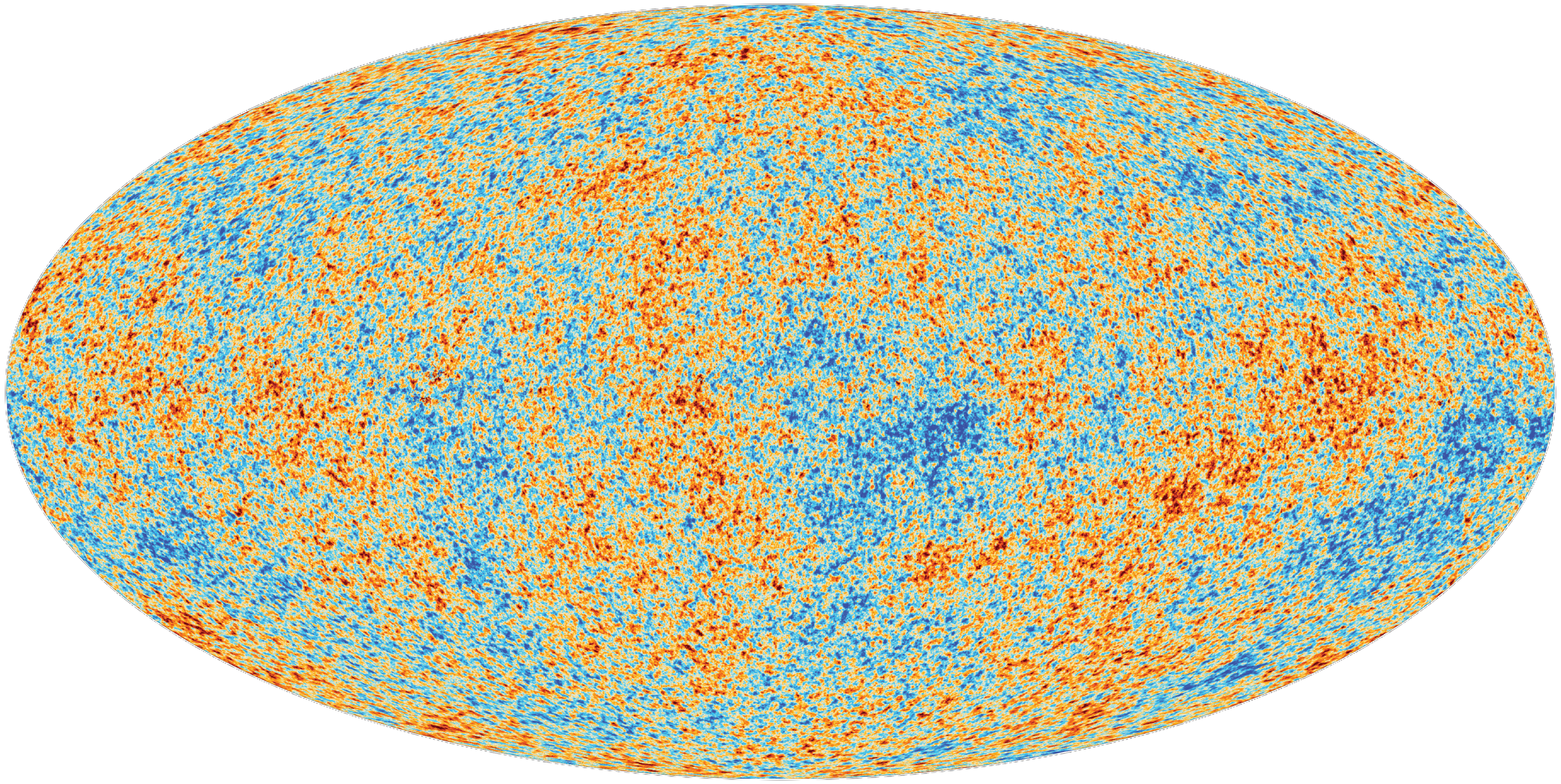
4. Build an experiment to get the data

5. Produce many versions of what the data would look like for different underlying models and parameters

6. Run statistical analysis to find which model provides the best fit to the data

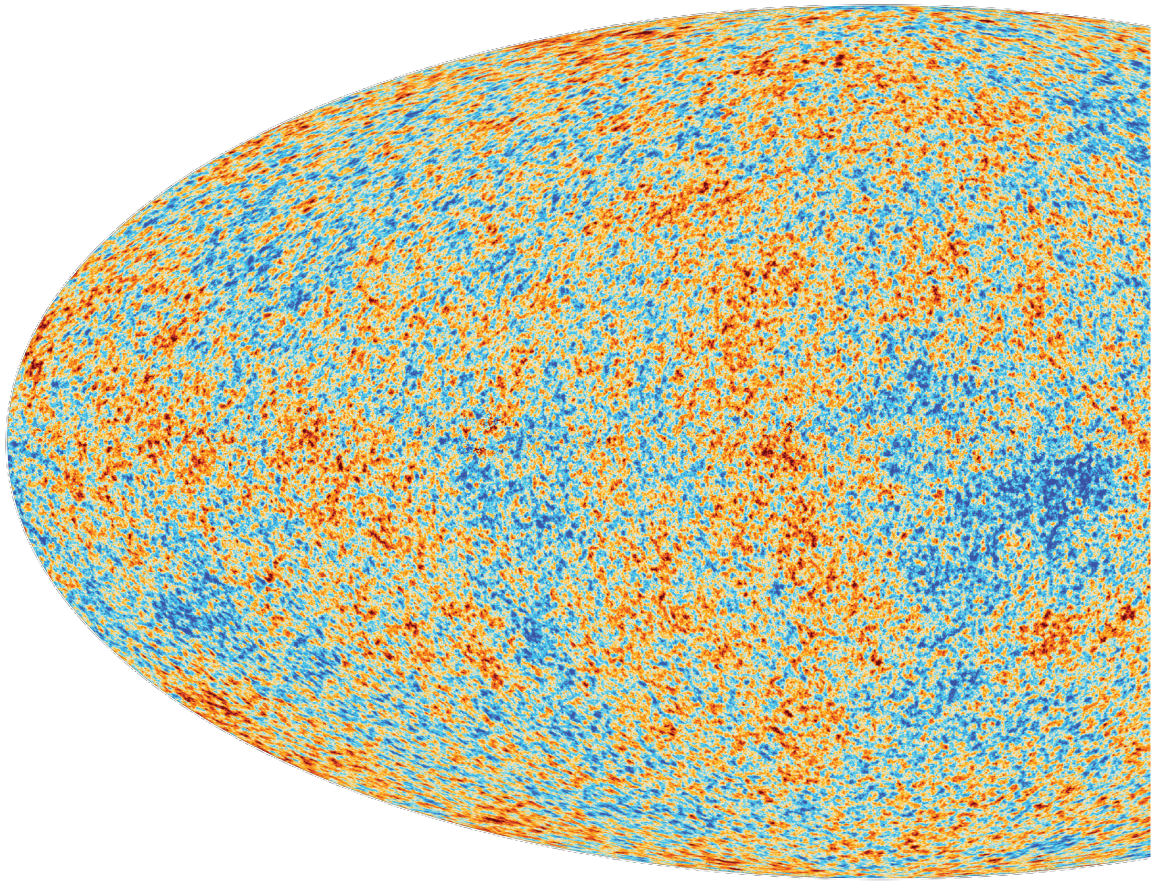
What type of equations go into our theories

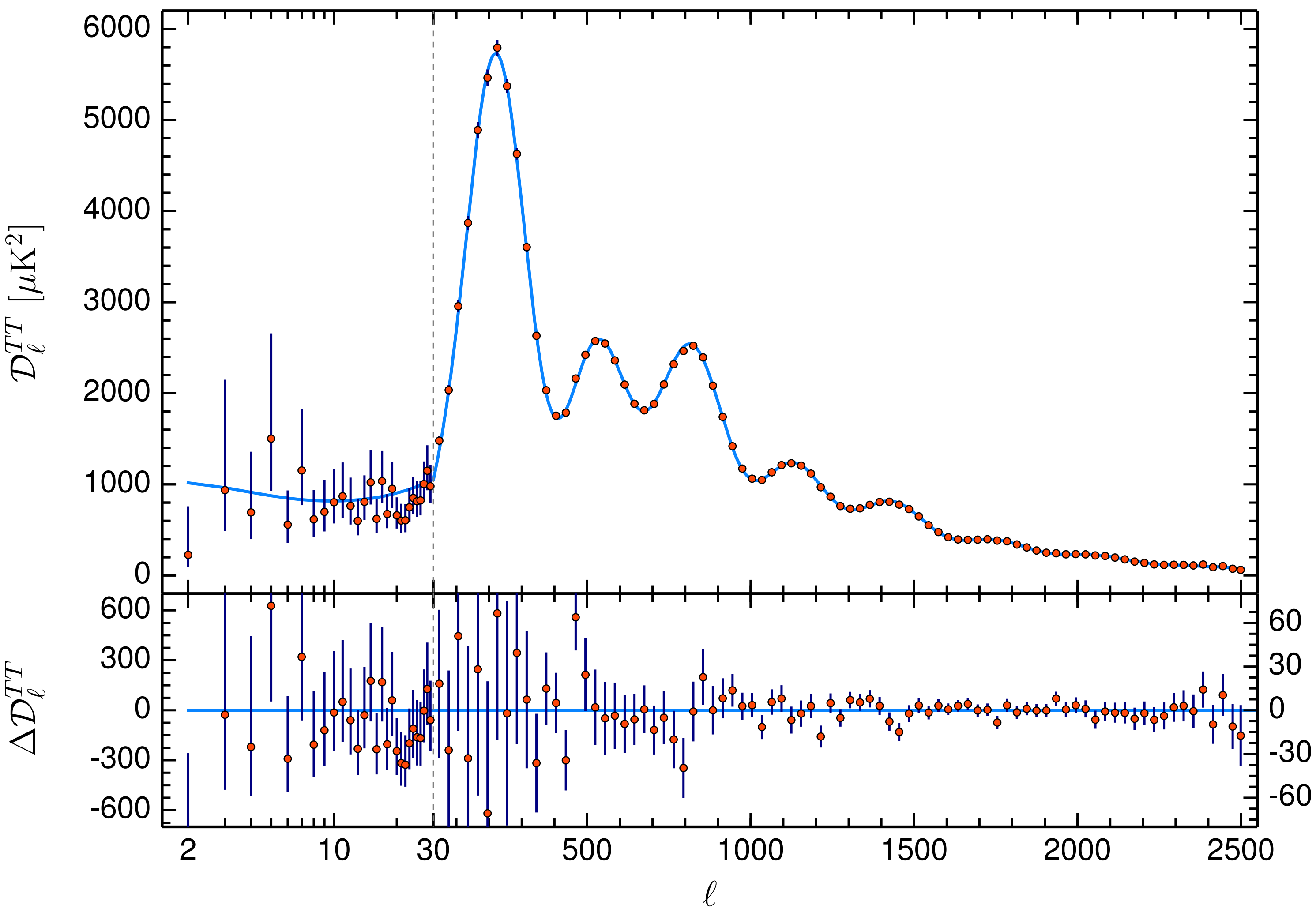
The CMB map gives us the temperature anisotropies power spectrum



Figures by Planck Collaboration

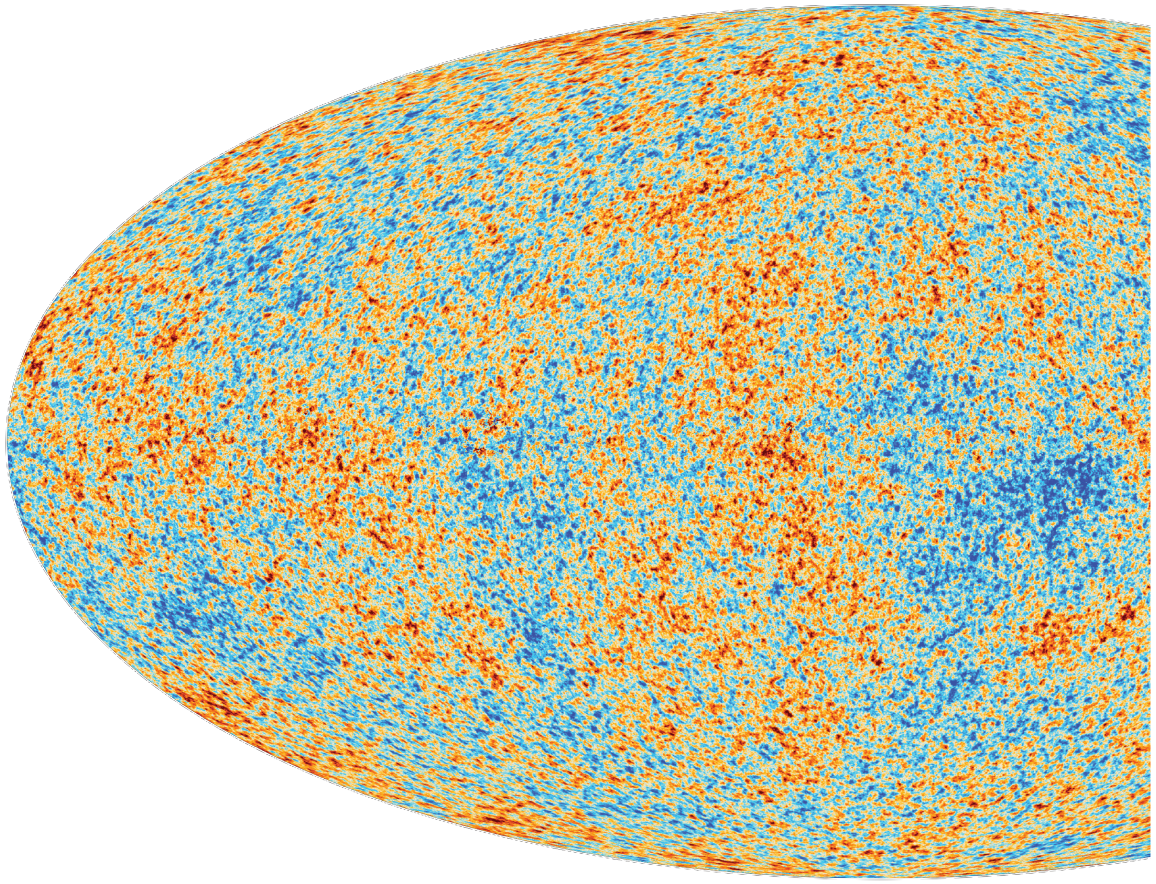
The CMB map gives us the temperature anisotropies power spectrum

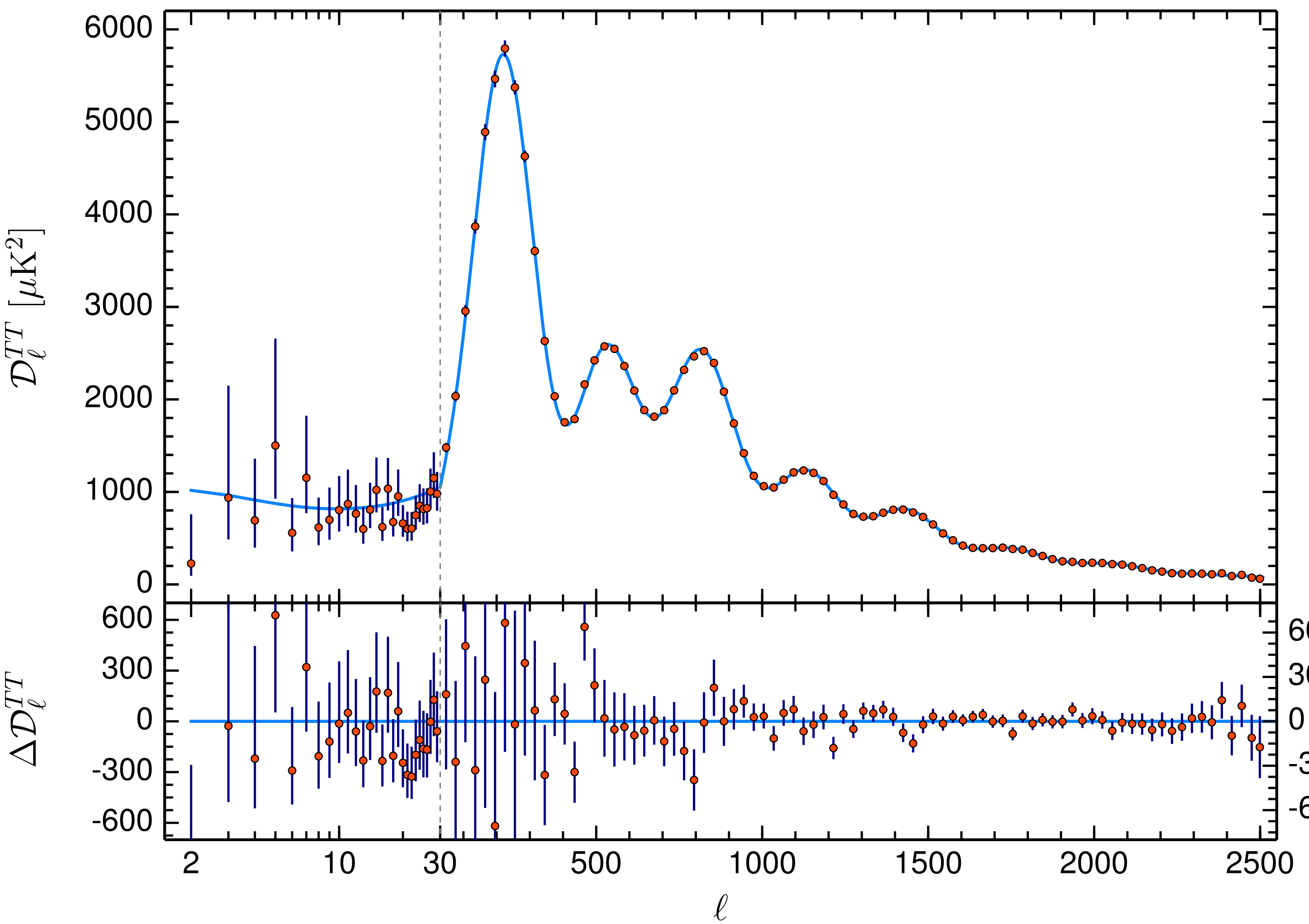




Figures by Planck Collaboration

The CMB map gives us the temperature anisotropies power spectrum





$$\Theta_{T,\ell}(\tau_0,k) = \int_{\tau_{\rm ini}}^{\tau_0} d\tau S_T(\tau,k) j_\ell(k(\tau_0-\tau))$$

$$S_T \equiv \underbrace{g(\Theta_0+\psi)}_{\text{SW}} + \underbrace{\frac{d}{d\tau}\left(\frac{g}{k^2}\theta_b\right)}_{\text{Doppler}} + \underbrace{e^{-\kappa}\left(\dot\phi+\dot\psi\right)}_{\text{ISW}} - \underbrace{\left(\frac{3}{4k^2}\frac{d^2}{d\tau^2}(\Pi g)+\frac{1}{4}\Pi g\right)}_{\text{Polarisation}}$$

Figures by Planck Collaboration

Observations of galaxies and lensing gives us the matter power spectrum

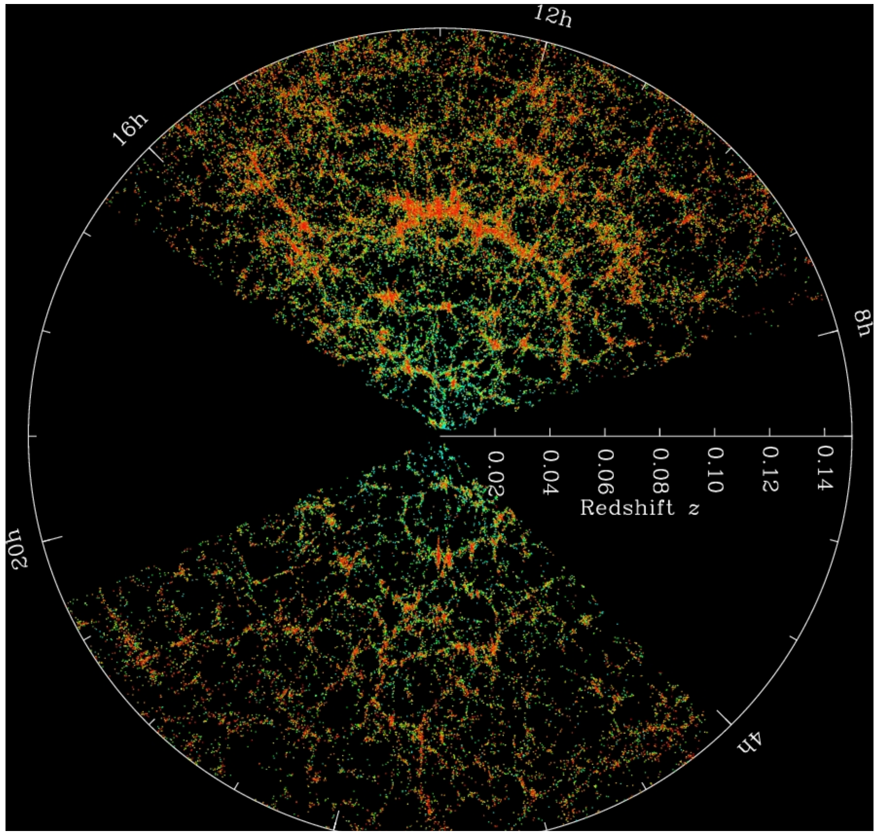
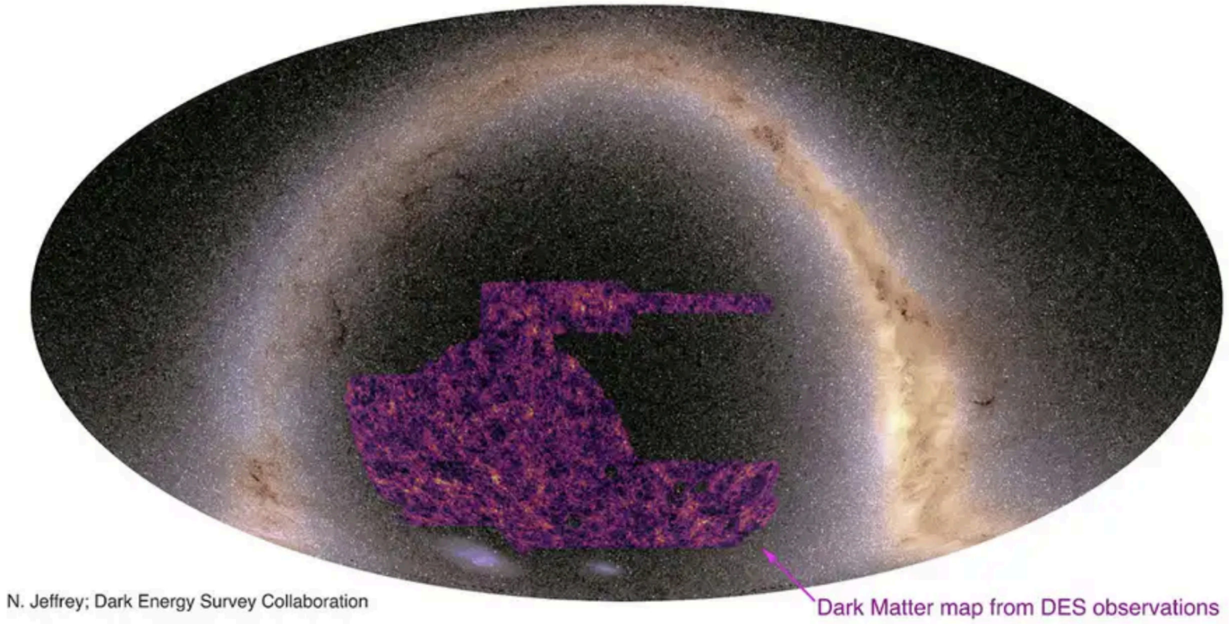


Figure by SDSS



Observations of galaxies and lensing gives us the matter power spectrum

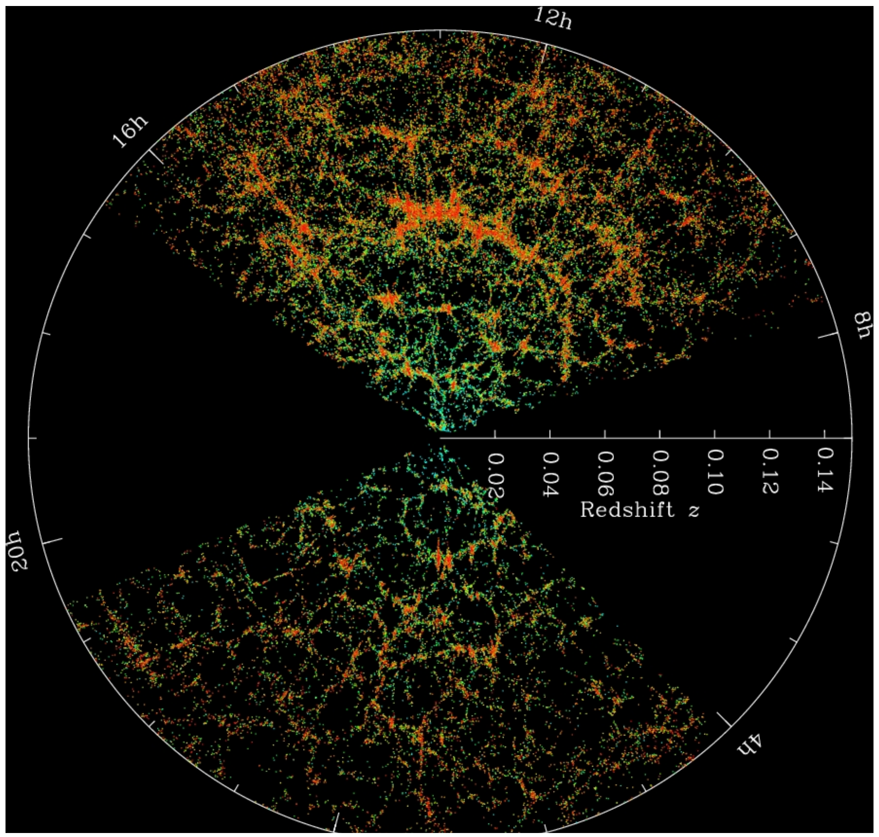
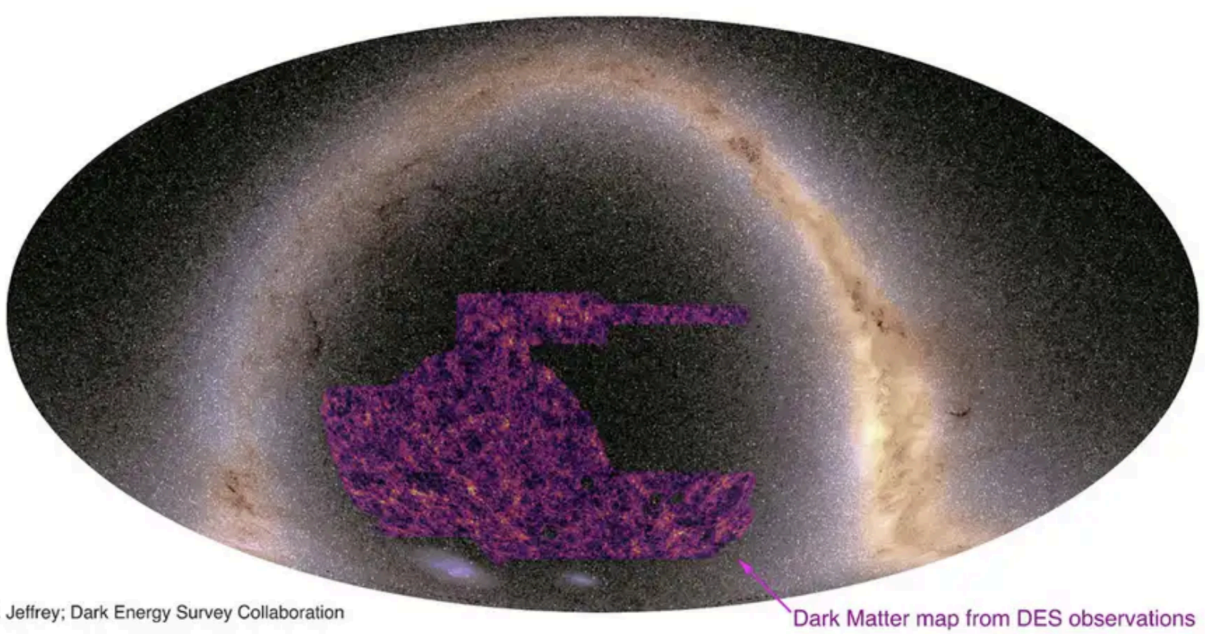


Figure by SDSS



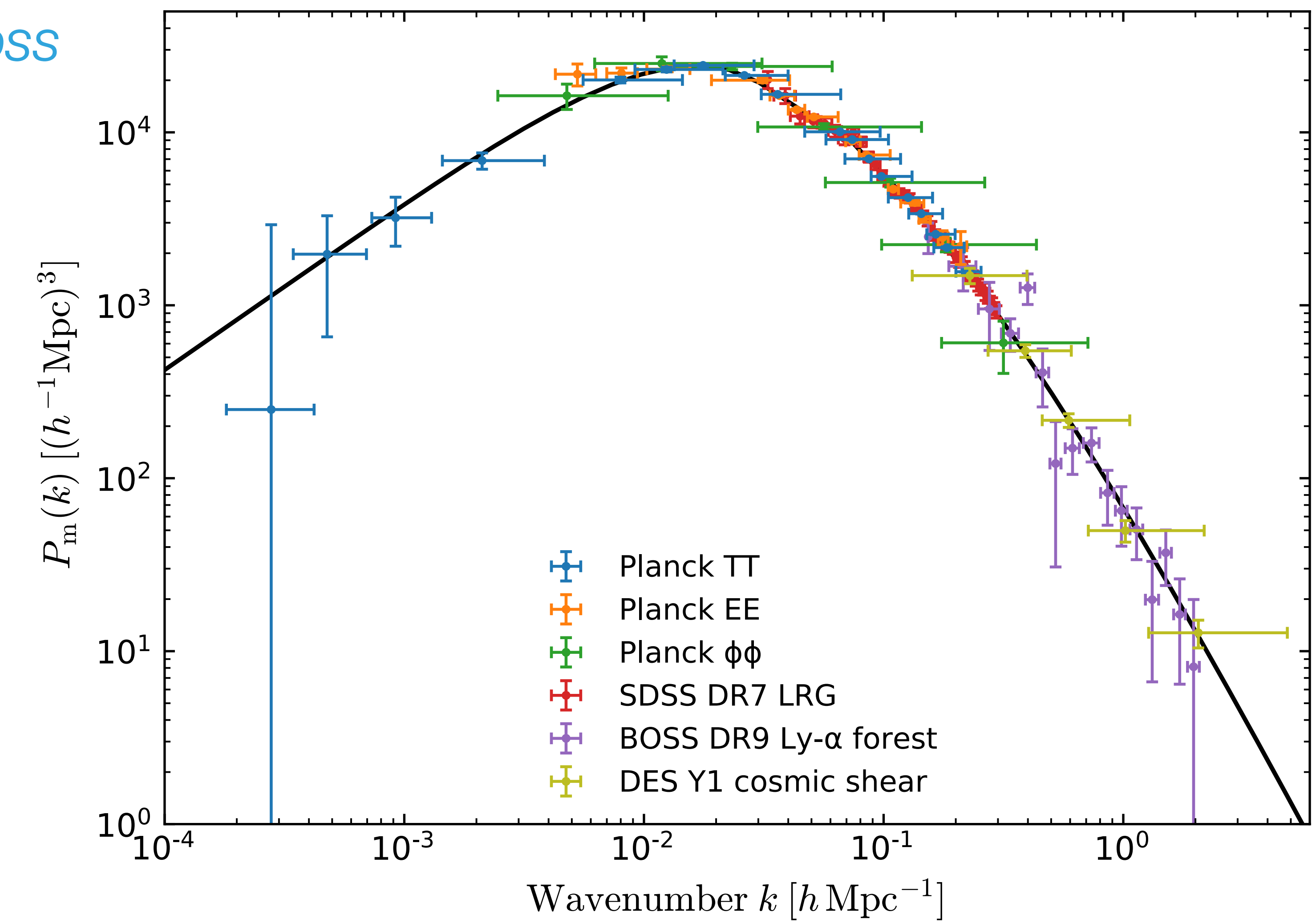


Figure by Planck Collaboration

Observations of galaxies and lensing gives us the matter power spectrum

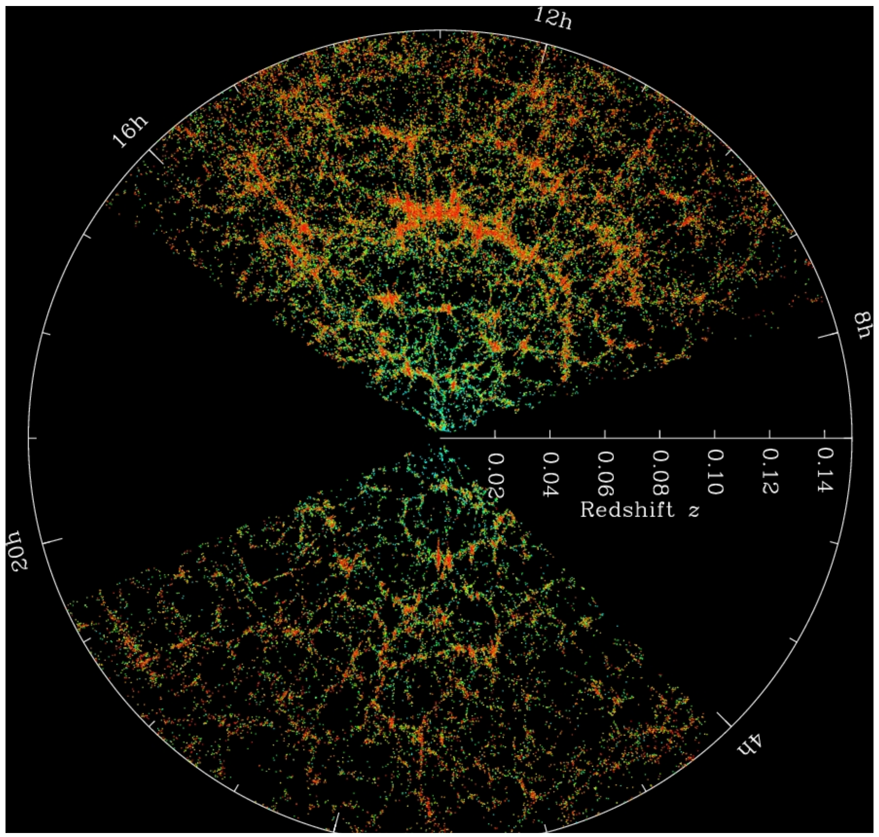
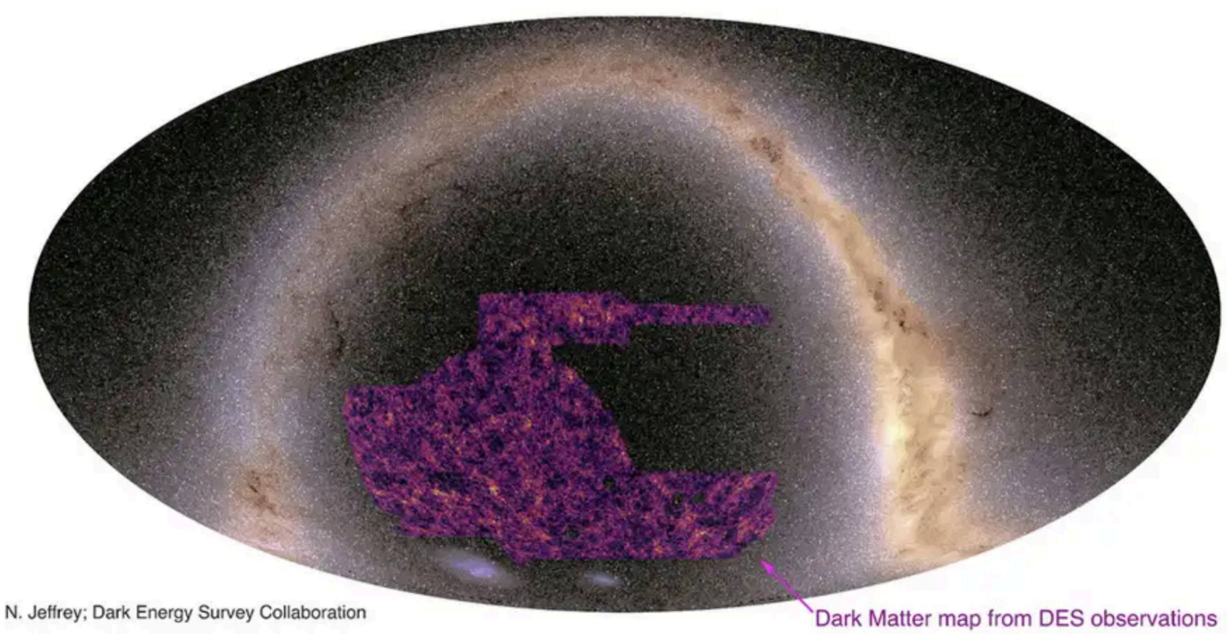


Figure by SDSS



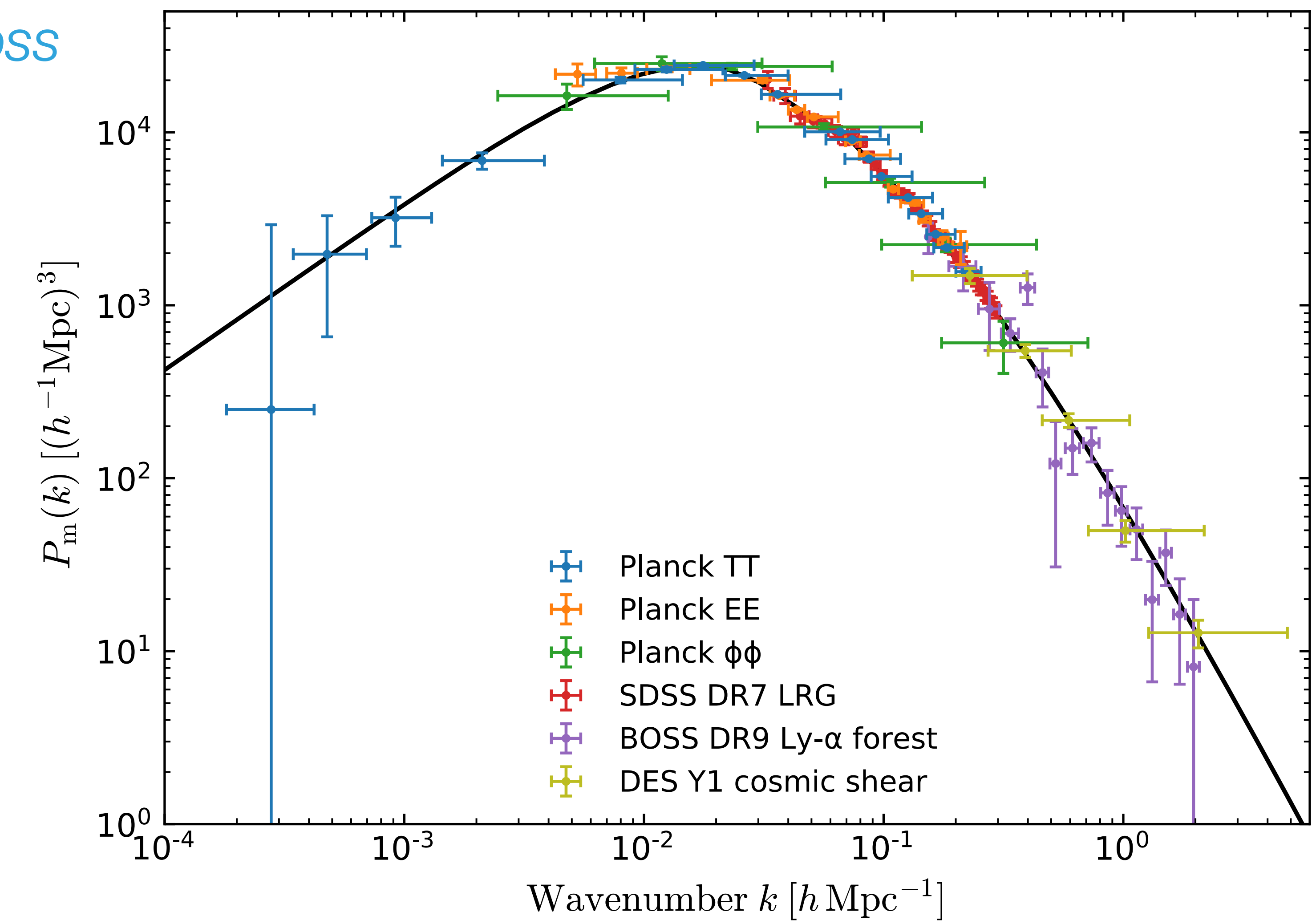


Figure by Planck Collaboration

$$\ddot{\delta}_{\rm CDM} + \frac{\dot{a}}{a}\dot{\delta}_{\rm CDM} - 4\pi G a^2 \bar{\rho}_{\rm CDM}\delta_{\rm CDM} = 0$$

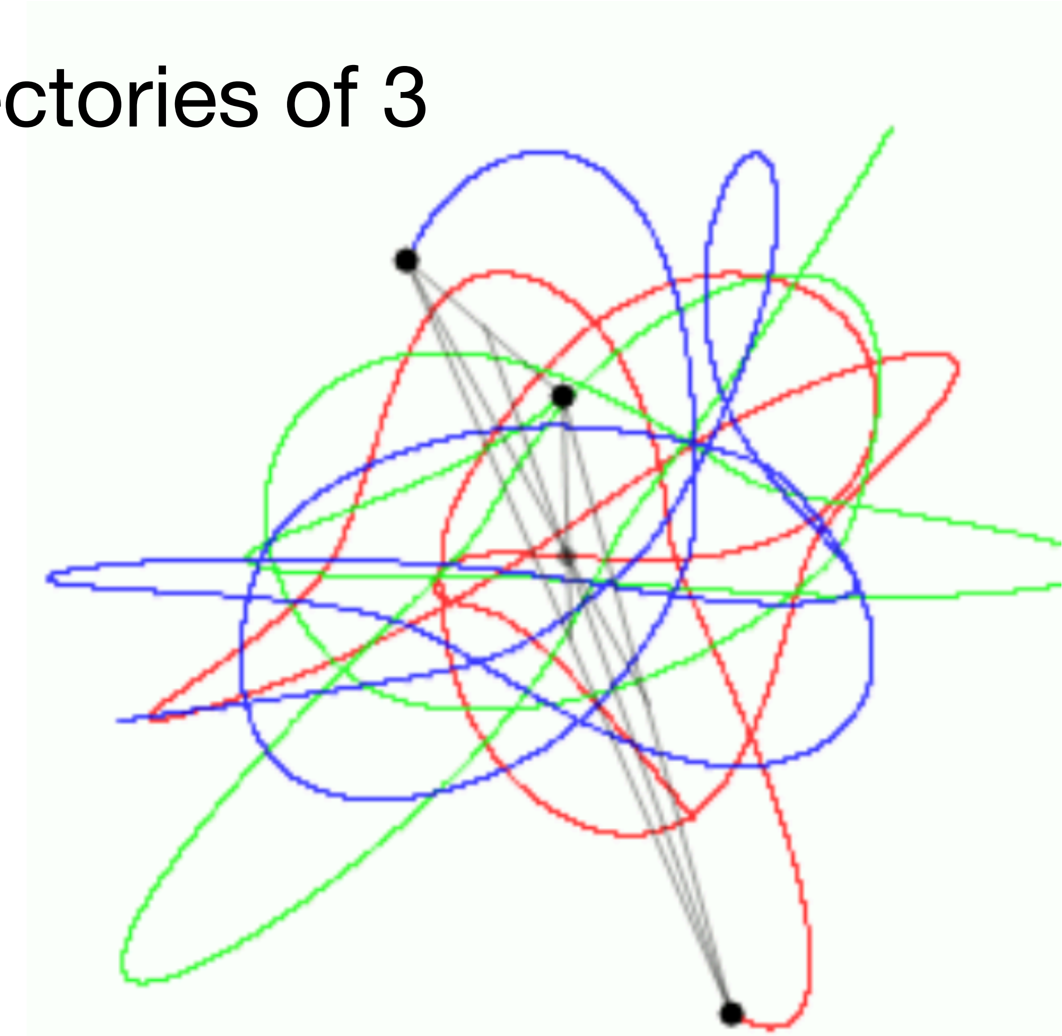
The three body problem cannot be solved analytically for all initial conditions

- For arbitrary initial conditions, trajectories of 3 gravitationally interacting bodies:

$$\ddot{\mathbf{r}}_1 = -Gm_2\frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} - Gm_3\frac{\mathbf{r}_1 - \mathbf{r}_3}{|\mathbf{r}_1 - \mathbf{r}_3|^3}$$

$$\ddot{\mathbf{r}}_2 = -Gm_3\frac{\mathbf{r}_2 - \mathbf{r}_3}{|\mathbf{r}_2 - \mathbf{r}_3|^3} - Gm_1\frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_2 - \mathbf{r}_1|^3}$$

$$\ddot{\mathbf{r}}_3 = -Gm_1\frac{\mathbf{r}_3 - \mathbf{r}_1}{|\mathbf{r}_3 - \mathbf{r}_1|^3} - Gm_2\frac{\mathbf{r}_3 - \mathbf{r}_2}{|\mathbf{r}_3 - \mathbf{r}_2|^3}$$



Animation available here

- When modelling galaxies or clusters, we need to generalise this to n bodies

How to put equations on a computer

Numerically solving the 3-body problem: Euler approach

1. Reduce differential equations to first order

$$\frac{dv}{dt} = a(t) \quad \& \quad \frac{dx}{dt} = v(t)$$

2. Numerical solution from Taylor expansion

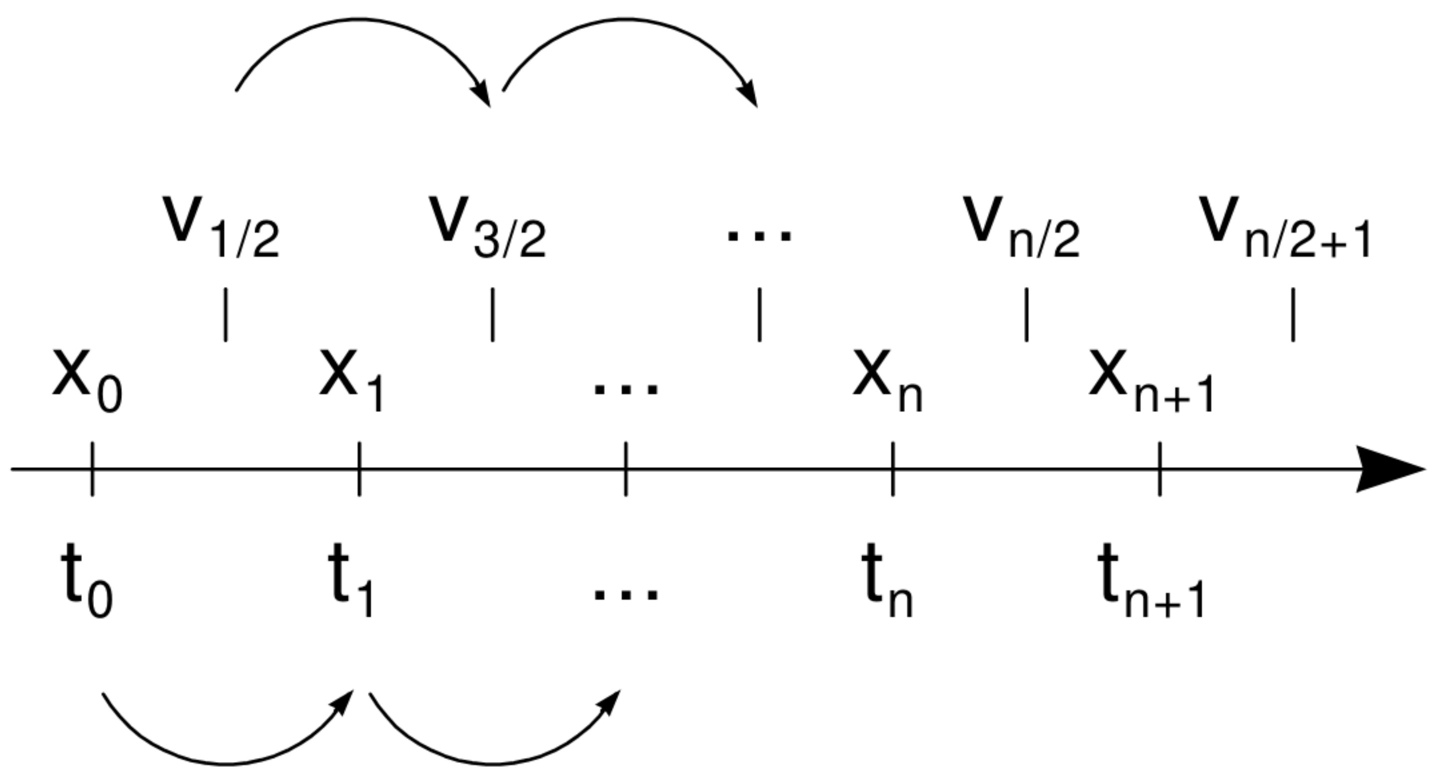
$$\begin{aligned} v_{n+1} &= v_n + a_n \Delta t + \mathcal{O}((\Delta t)^2) \\ x_{n+1} &= x_n + v_n \Delta t + \frac{1}{2} a_n (\Delta t)^2 + \mathcal{O}((\Delta t)^3) \end{aligned}$$

3. Euler method: account only for $\mathcal{O}(\Delta t)$ (for $\Delta t \to 0$)

$$\begin{aligned} v_{n+1} &= v_n + a_n \Delta t \\ x_{n+1} &= x_n + v_n \Delta t \end{aligned}$$

Numerically solving the 3-body problem: Leapfrog approach

- Leapfrog integration: lower error, same number of steps as Euler time-reversible

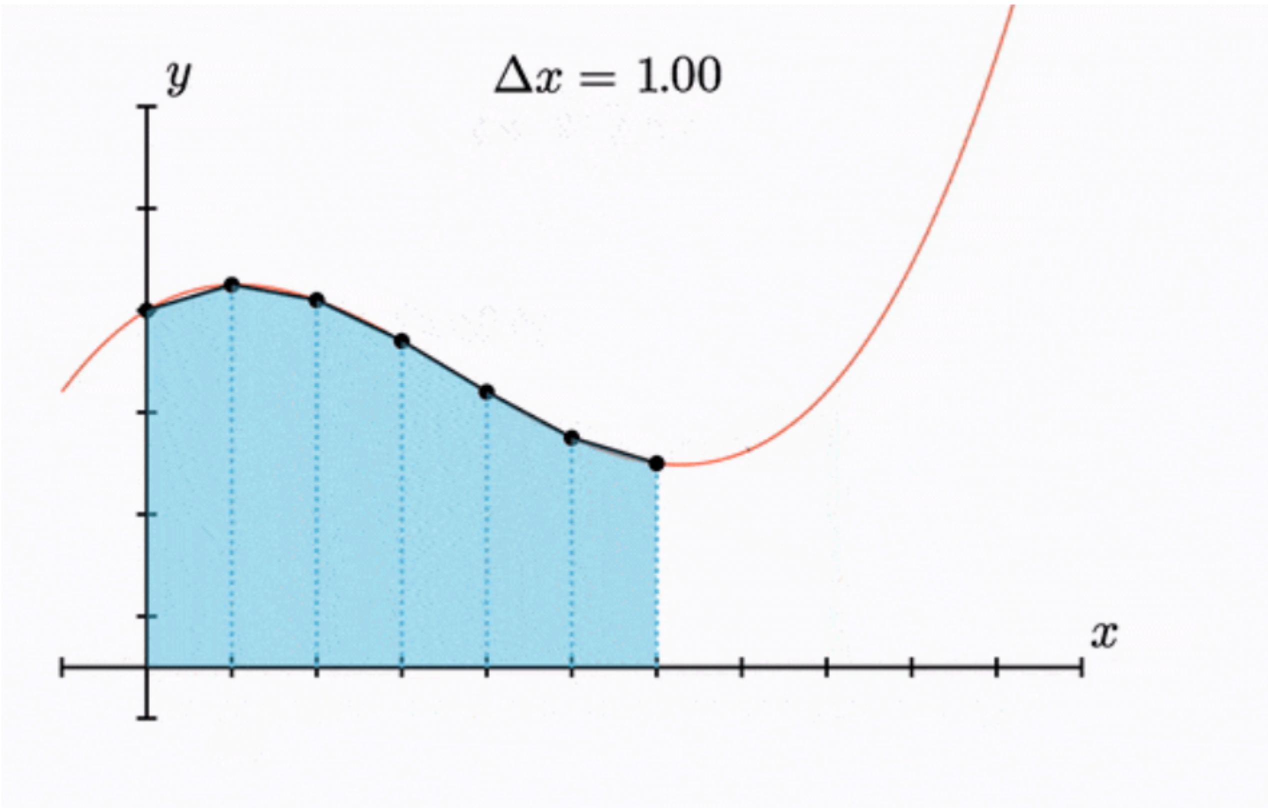


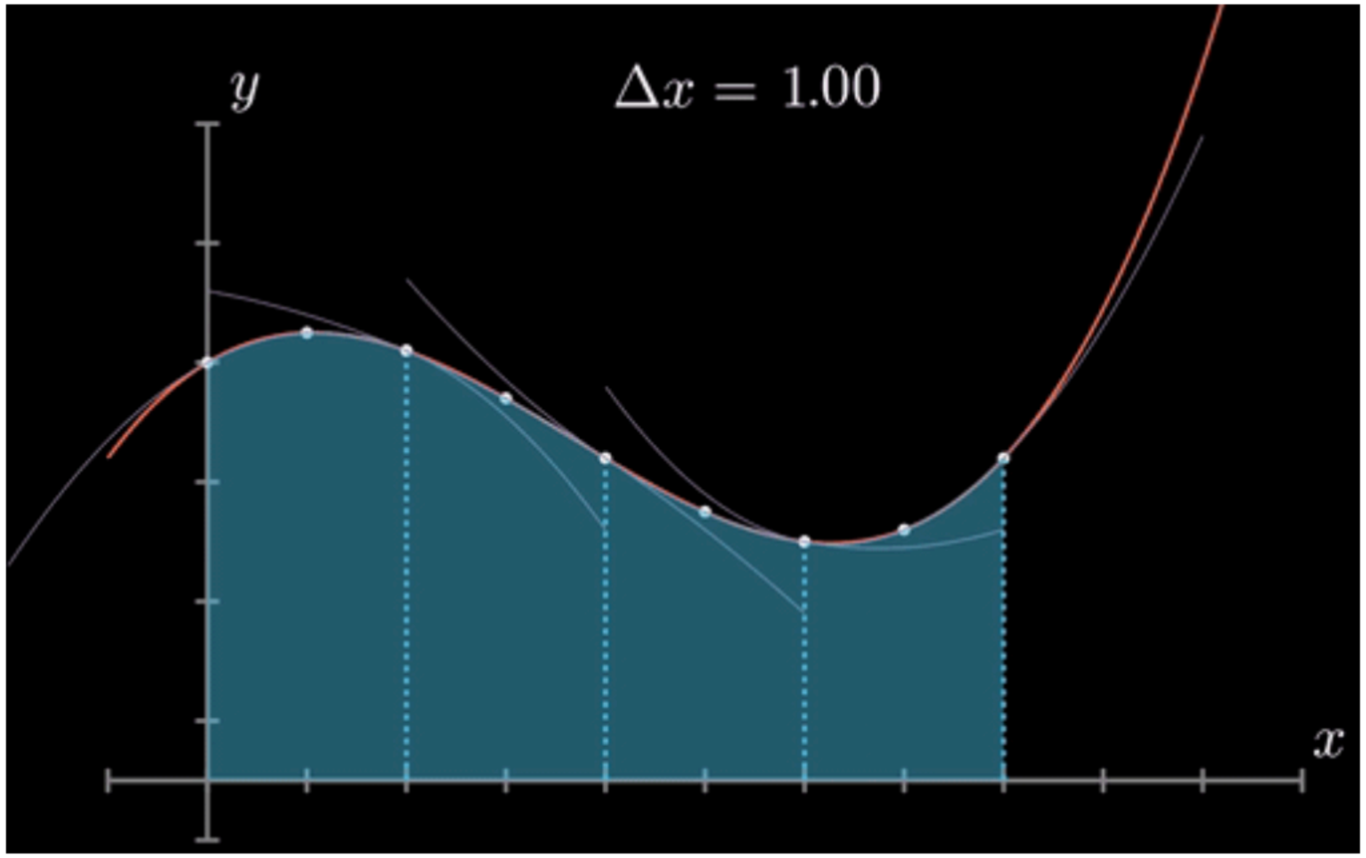
$$v_{n+\frac{1}{2}} = v_{n-\frac{1}{2}} + a_n \Delta t$$
$$x_{n+1} = x_n + v_{n+\frac{1}{2}} \Delta t$$

- Not self starting, so we use Euler for first step

$$v_{\frac{1}{2}} = v_0 + \frac{1}{2} a_0 \Delta t$$

To solve an integral, you can use things like Trapezoidal or Simpson's rule





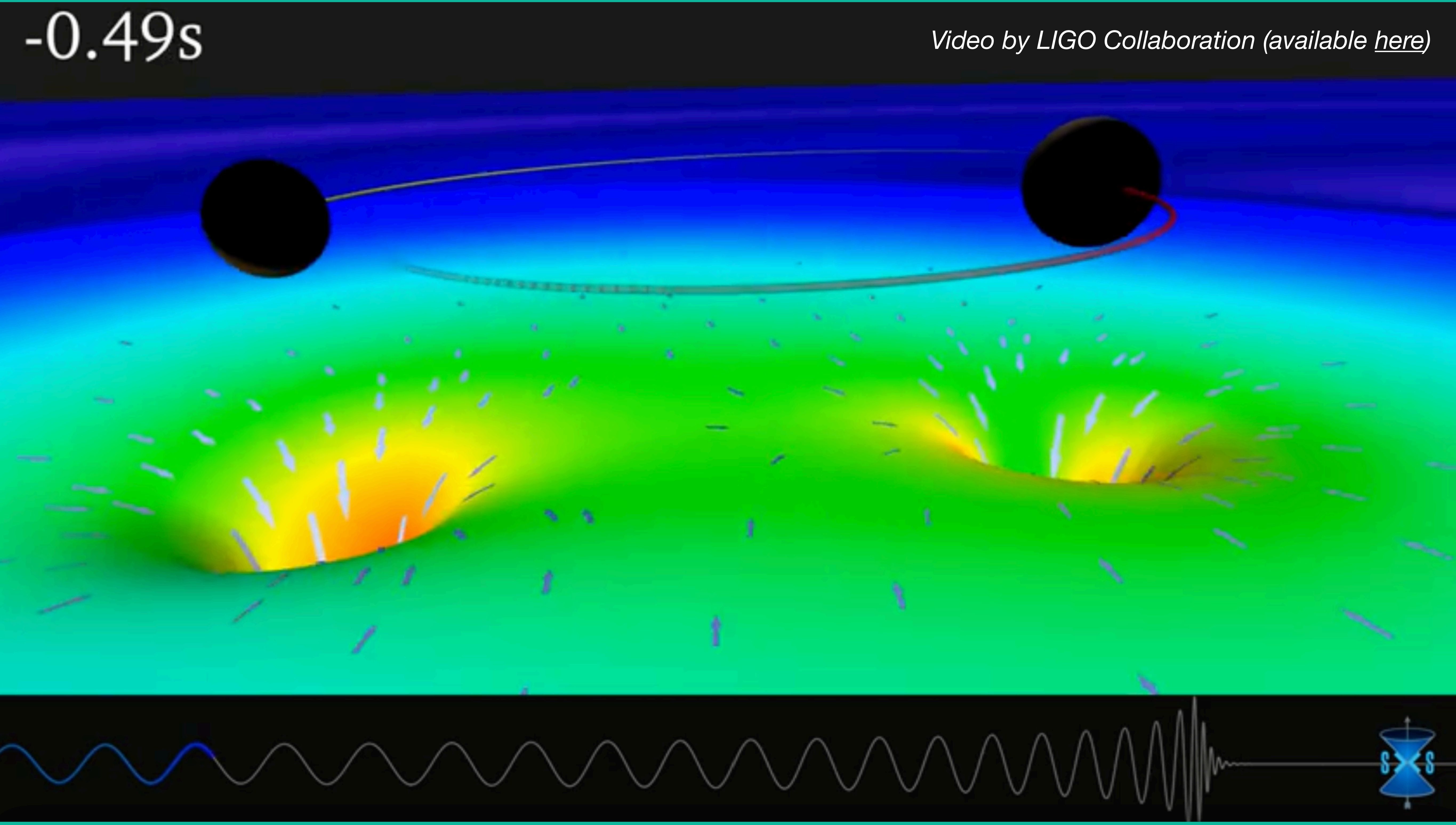
- Both methods rely on dividing the function and calculating the area under the curve
- Simpson's rule gives lower errors
- For narrow peak-like functions trapezoidal rule might be more efficient

Animations by Aravindh Vasu (available here and here)

How to get the computer to do math for you: simplify!

- Can you break the problem down into smaller steps?
- Can you split it into different parameter regions?
- Are there any singularities you can remove?
- Are there already established methods to solve these things? How can you adapt that to your needs?
- You will often have a trade-off between precision and computational speed

Now you can make cool simulations!



Video by LIGO Collaboration (available here)

Model comparison and statistics

How do you compare different models that seem to describe the data well?

- For a model *M* and observed data *D,* Pearson's χ^2 test:

$$\chi^2 = \sum_i \frac{(D_i - M_i)^2}{D_i}$$

- The model that gives the lowest χ^2 provides the best fit to the data

- This assumes that the different parameters follow gaussian distributions, and that there are no parameter correlations

- For a more thorough analysis, we use likelihood functions

The likelihood function tells you how likely it is that a model describes the data

- For a model *M* described by parameters θ and observed data *D*
- The likelihood function ($\mathscr{L}$) tells you how likely an observed *D* is for a given set of θ belonging to *M*
- Maximise likelihood (minimise the χ^2): find which set of parameters θ best describe the data *D*
- Can also be used to compare different models *M*

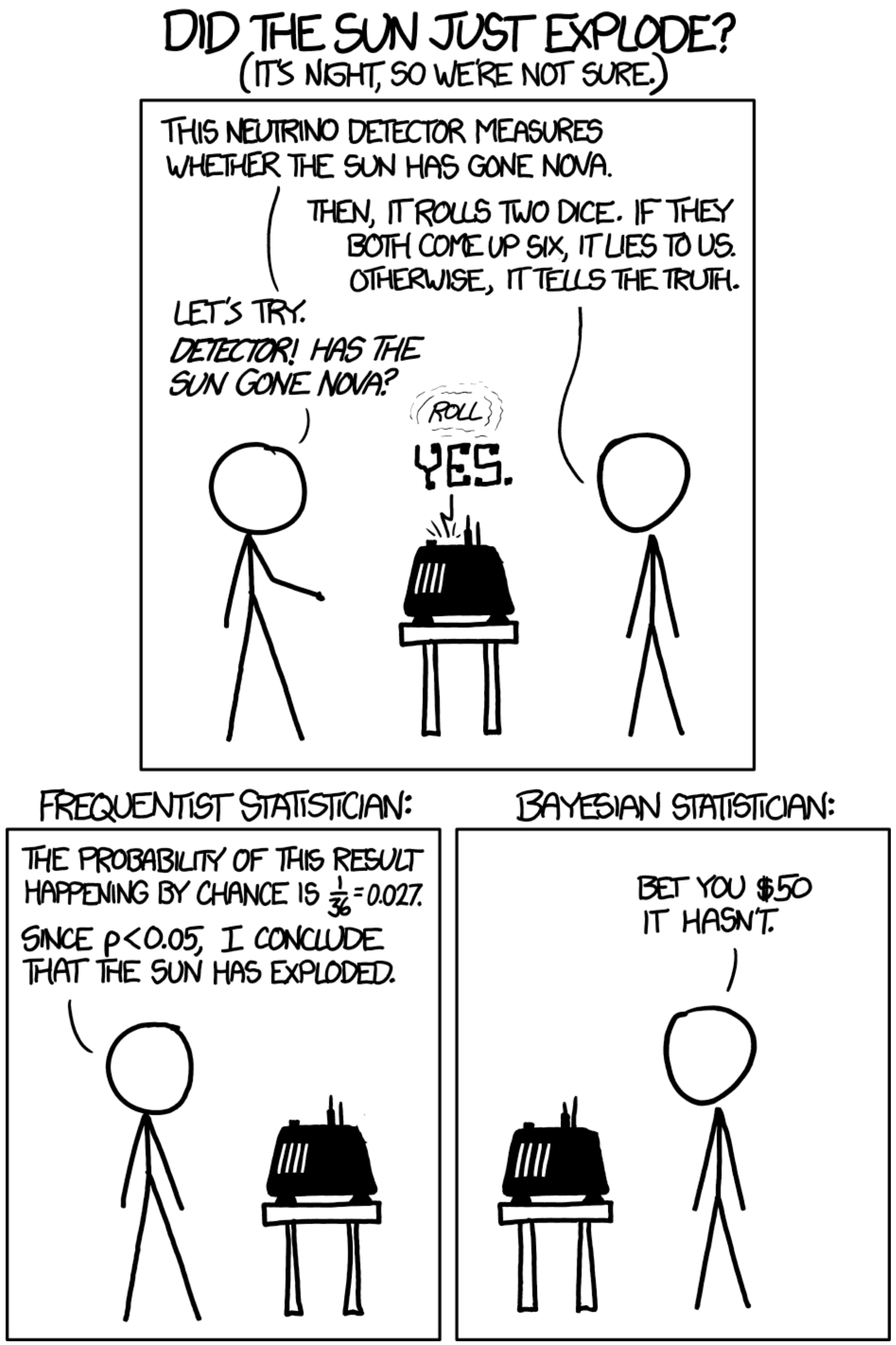
Likelihoods and probability sound similar, but they answer different questions

- Probability gives chance of D for a given M and θ
- Likelihood takes D as the starting point, and sees if this can be explained by M and θ
- In a fair coin, you have a 0.5 probability of getting heads
- Likelihood: flip coin 50 times, get 10 heads, how likely is it that coin is fair?

Markov Chain Monte Carlo: steps in parameter space to see how likely each step is

- Start at a point in parameter space with $\theta_0, \mathscr{L}_0$
- Take a random step to a new point $\theta_1, \mathscr{L}_1$
- If $\mathscr{L}_1 > \mathscr{L}_0$ accept the point, otherwise reject it most of the time, but allow for occasional steps to a lower likelihood
- Keep doing this until you have a chain, create probability distribution function from the chain

How much prior knowledge should you include in your statistical tests?



- Do the priors matter?
- Flip coin 50 times, get 10 heads - do you still assume a fair coin?
- If priors affect the likelihood, you're doing Bayesian stuff

Figure by Randall Munroe, XKCD (1132)

From theories to observations: a step by step example

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

4. Build an experiment to get the data

5. Produce many versions of what the data would look like for different underlying models and parameters

6. Run statistical analysis to find which model provides the best fit to the data

From theories to observations: a step by step example

1. Come up with a theory or model that makes predictions

2. If your underlying equations cannot be solved analytically, you will need a computer to solve them numerically

3. Figure out which data let you test your predictions

- Model: ΛCDM
- Prediction: the shape of the temperature anisotropies power spectrum
- Use CLASS or CAMB to get the predicted power spectrum
- Data: the cosmic microwave background temperature anisotropies

From theories to observations: a step by step example

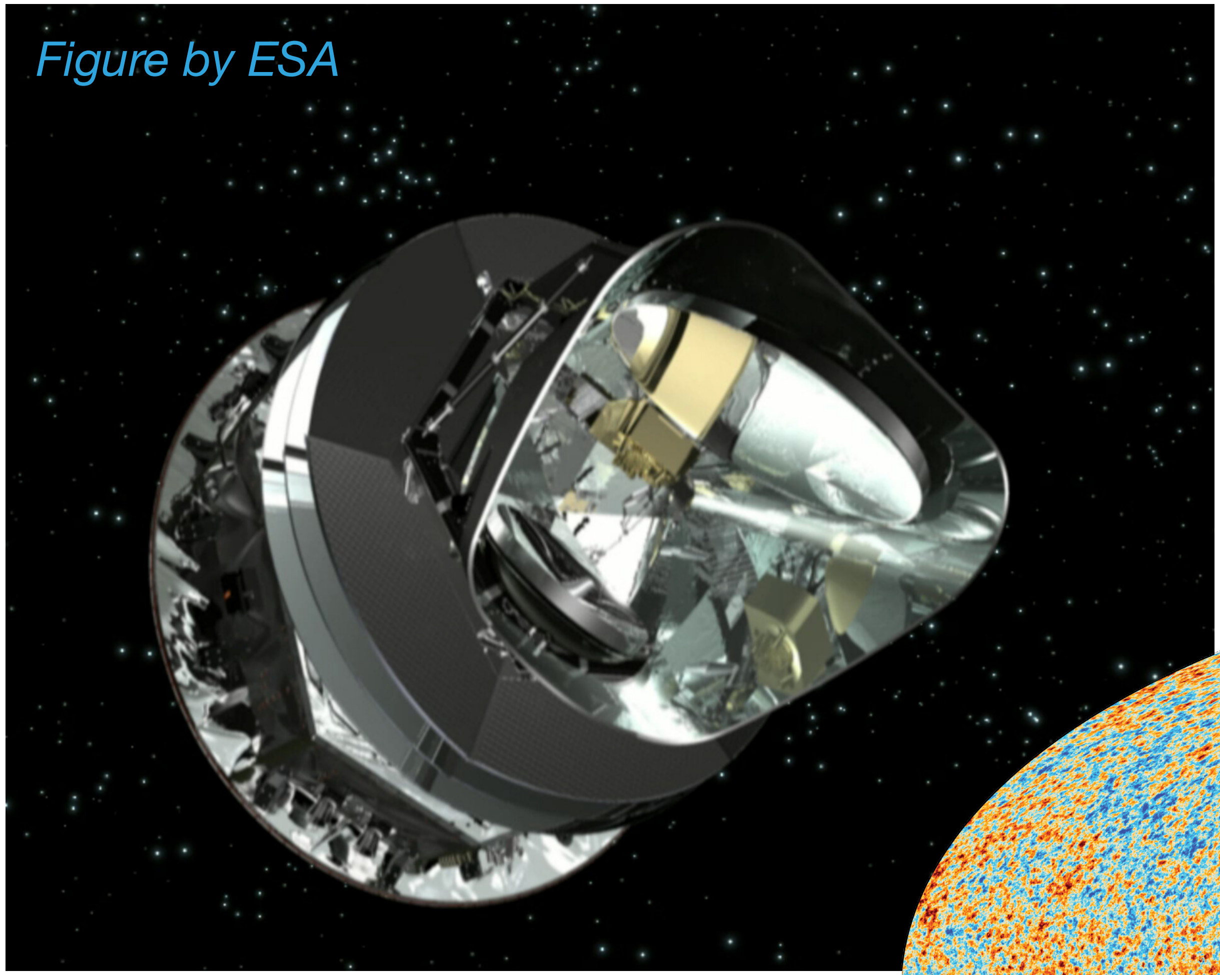


Figure by ESA

4. Build an experiment to get the data

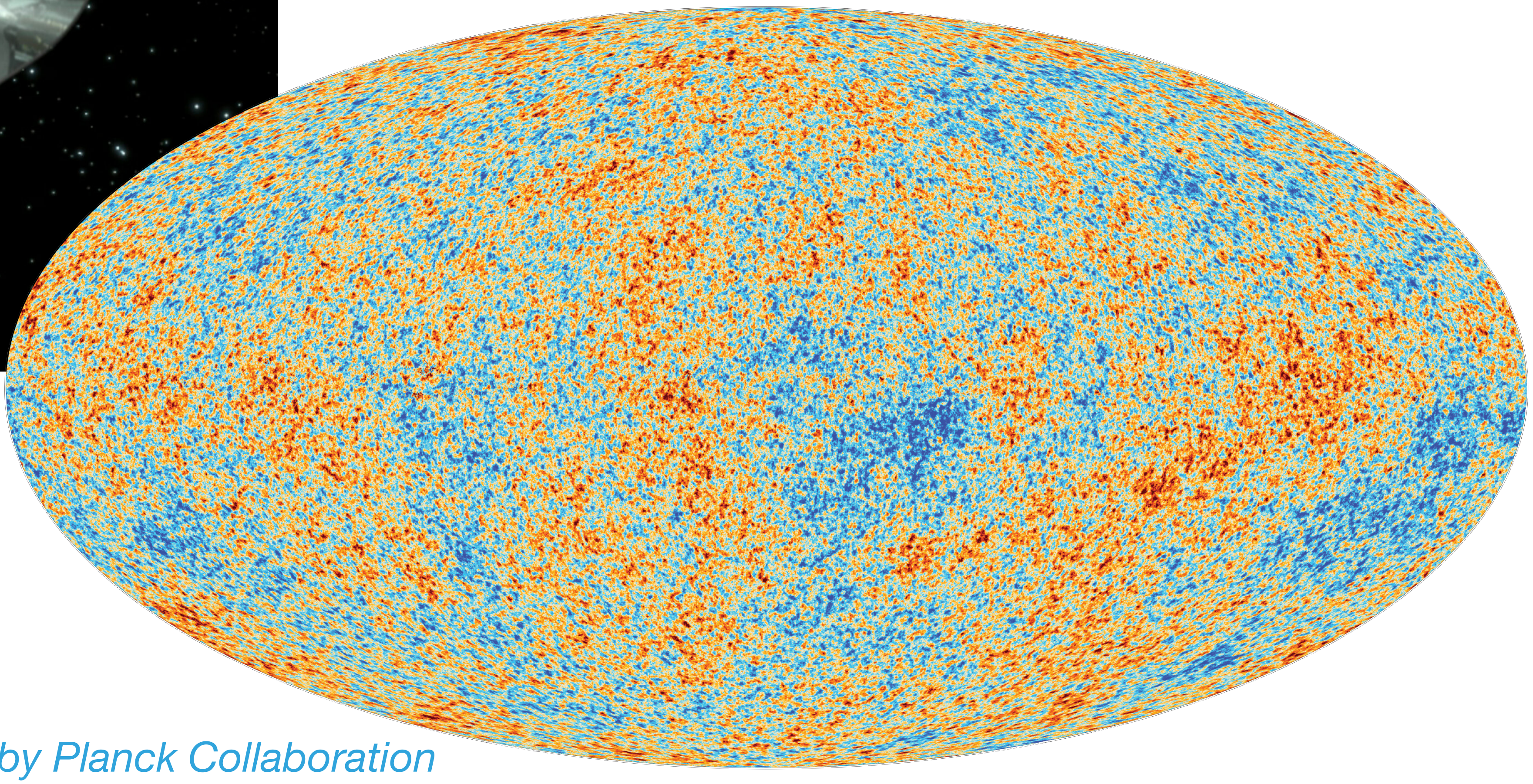
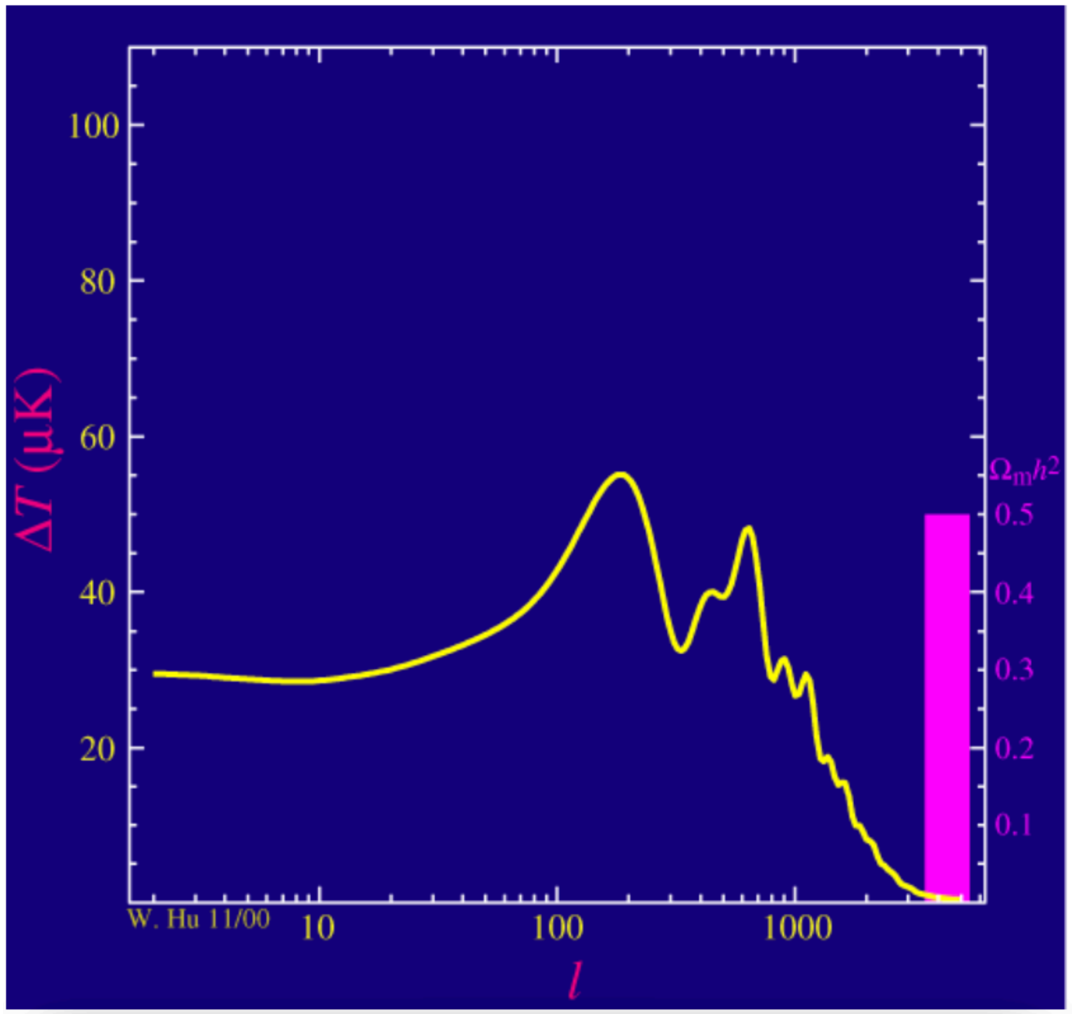


Figure by Planck Collaboration

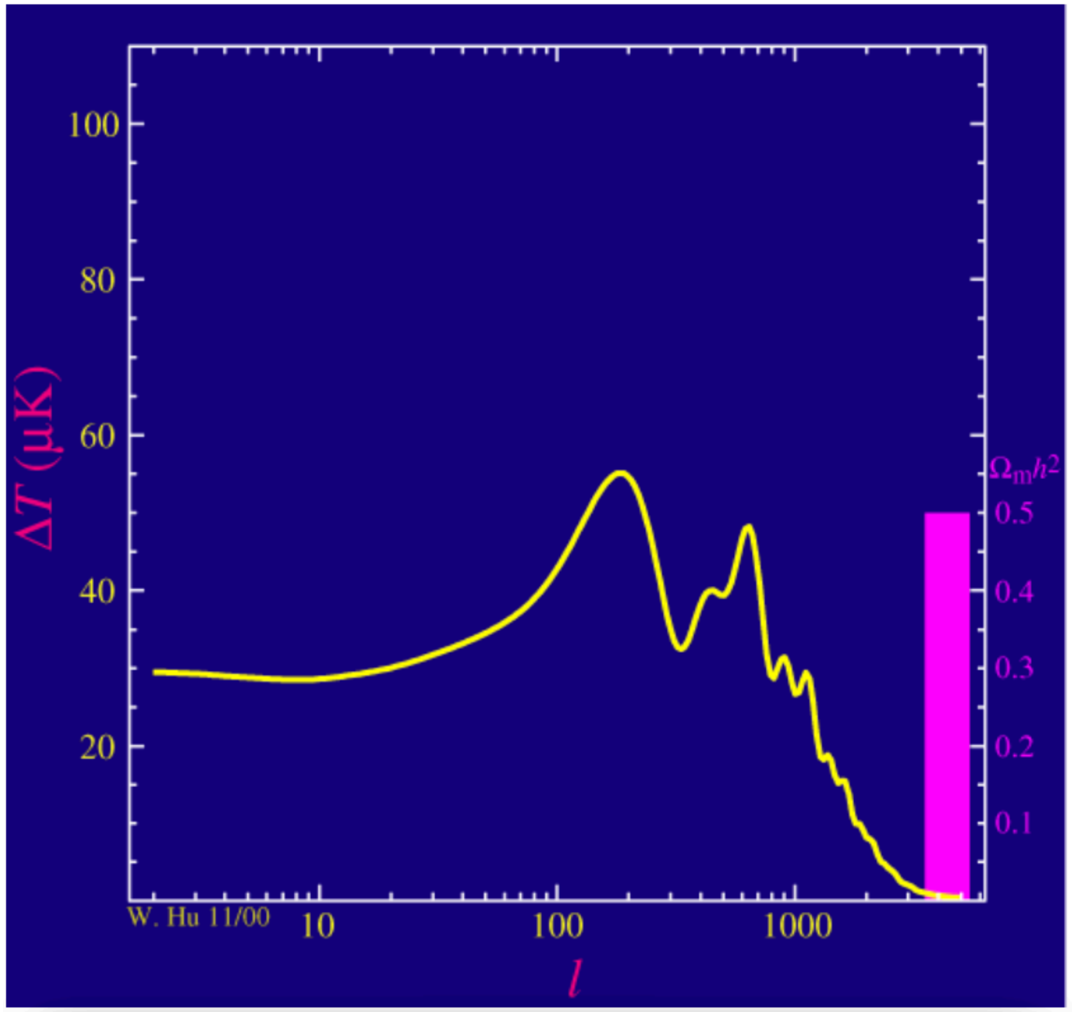
From theories to observations: a step by step example



Animation by Wayne Hu (available <u>here</u>)

5. Produce many versions of what the data would look like for different underlying models and parameters

From theories to observations: a step by step example

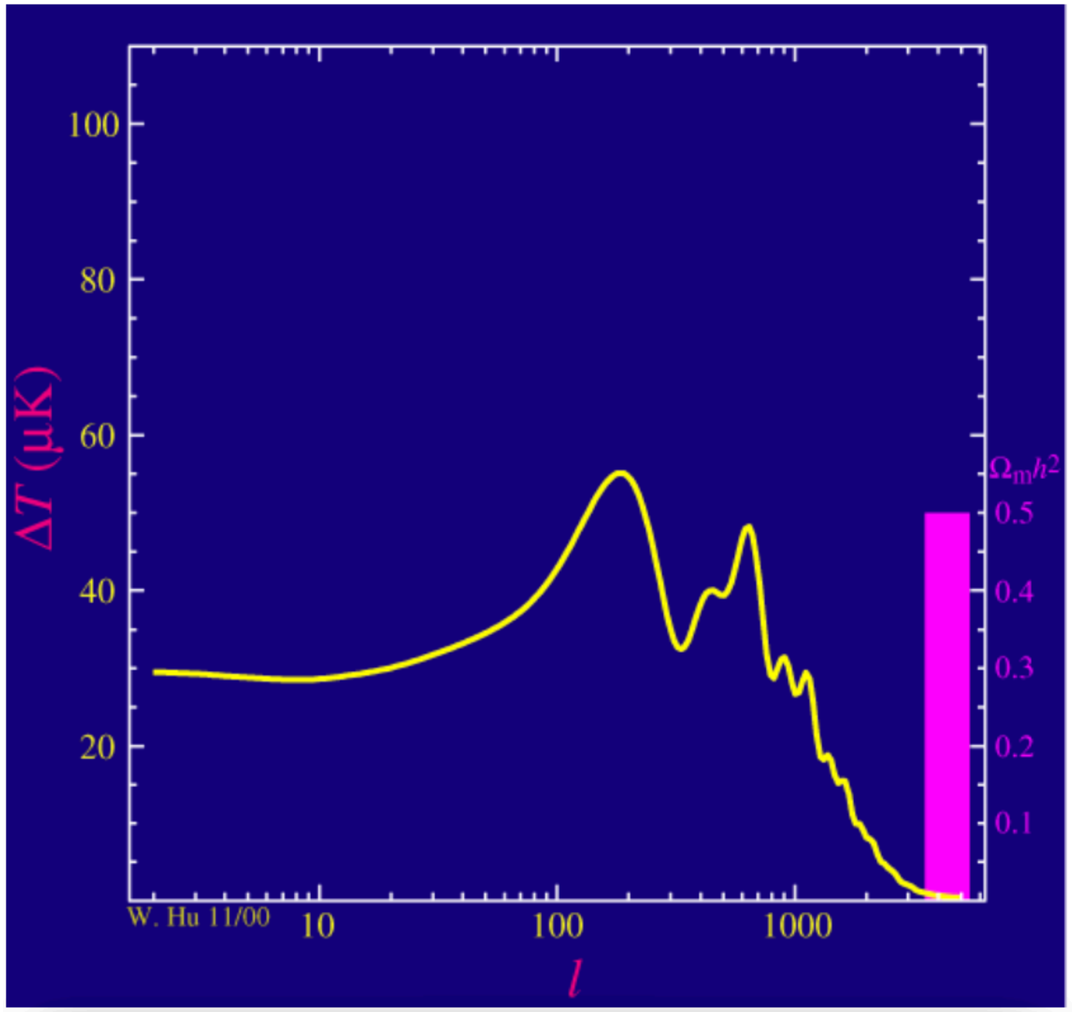


Animation by Wayne Hu (available here)

5. Produce many versions of what the data would look like for different underlying models and parameters

6. Run statistical analysis to find which model provides the best fit to the data

From theories to observations: a step by step example



Animation by Wayne Hu (available here)

5. Produce many versions of what the data would look like for different underlying models and parameters

6. Run statistical analysis to find which model provides the best fit to the data

- This is what you will do in the exercise session!

Thanks for listening!

Questions?

Email: deanna.hooper@helsinki.fi