Title: Cosmology Observations 1

Date: Jul 11, 2018 09:00 AM

URL: http://pirsa.org/18070010

Abstract:

Input: DVI - 1920x1080p@60Hz
Output: SDI - 1920x1080i@60Hz

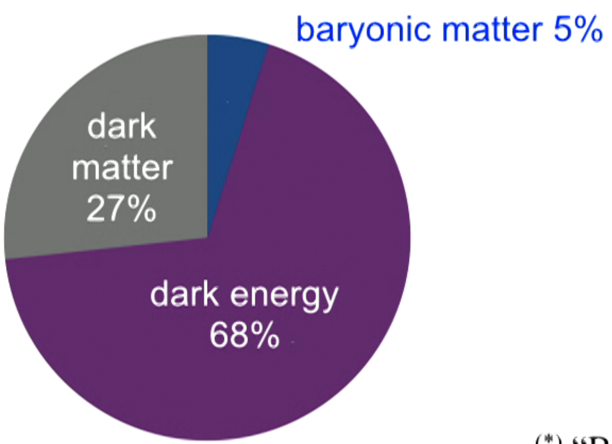
Observational cosmology

Kendrick Smith
TRISEP 2018

Part 1: the standard cosmological model

All current data can be fit by a 6-parameter cosmological model!

$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}\ \mathrm{GeV}^4$	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic(*) matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth



(*) "Baryons" = protons + neutrons + electrons(!)

Ingredients in the standard cosmological model:

- Background metric is FRW
- Expansion history is ΛCDM
- Initial perturbations are Gaussian random
- Initial perturbations are scalar adiabatic
- Power spectrum of initial perturbations is a power law: $(k^3/2\pi^2)P(k) = \Delta_\zeta^2 (k/k_0)^{n_s - 1}$

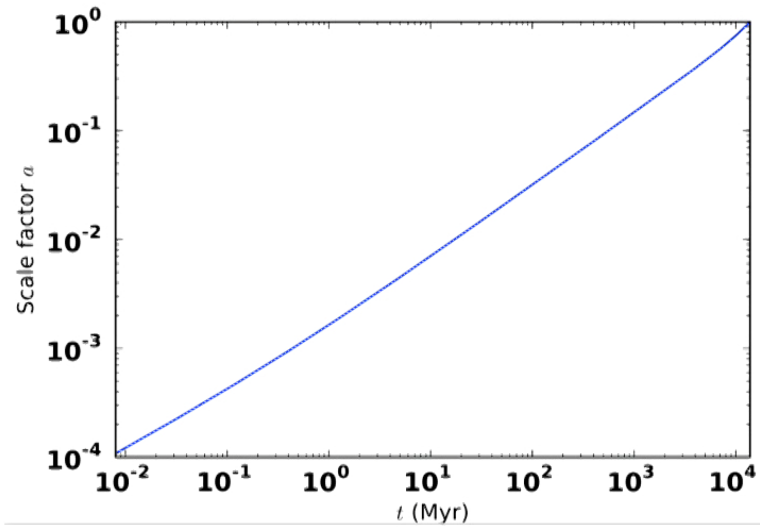
In the next few slides, we'll describe these ingredients at an informal level, just to set the stage. (Focus of these lectures is data analysis and statistics, theory lectures are next week!)

"Background metric is FRW"

The expansion of the universe is described by a function a(t), such that a=0 at the big bang, and a=1 today. (a = "scale factor")

Formal meaning: metric is $ds^2 = -dt^2 + a(t)^2\, dx^2$

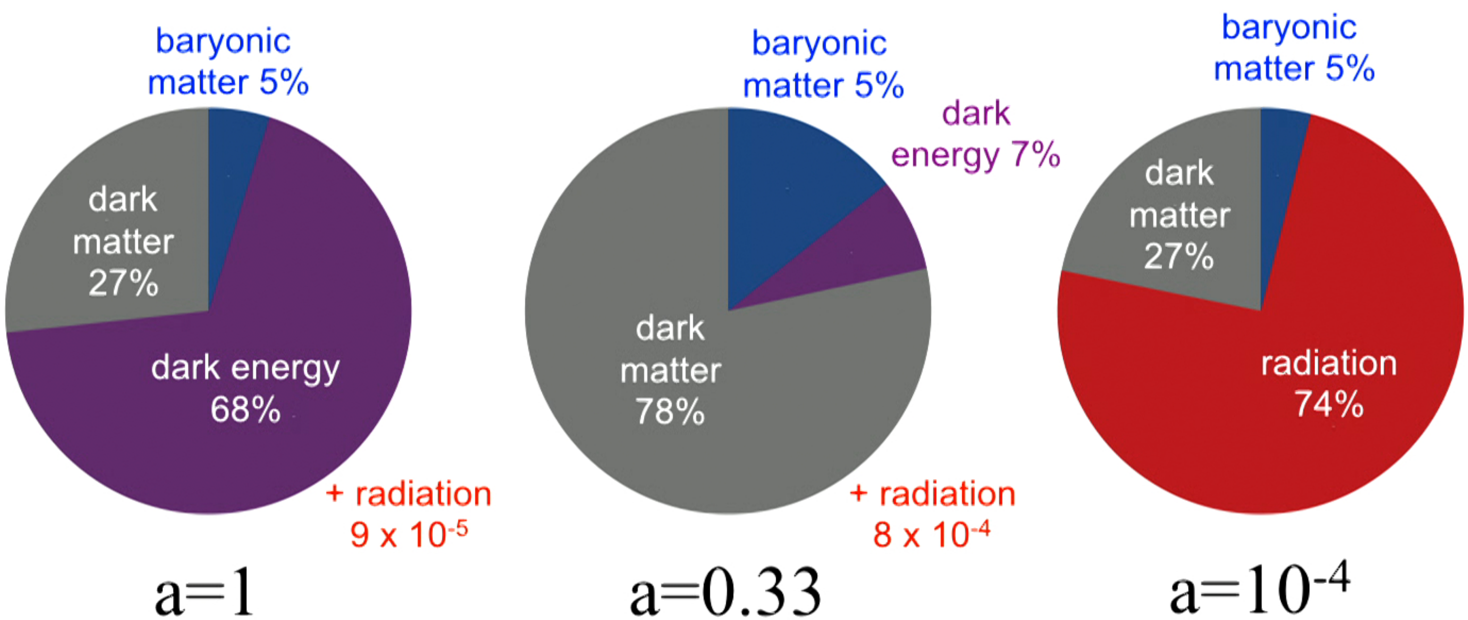
Intuitive meaning: if points x, x' are separated by distance D today, then their separation at time t is a(t)D.



“Expansion history is ΛCDM”

Energy densities evolve with scale factor a(t):

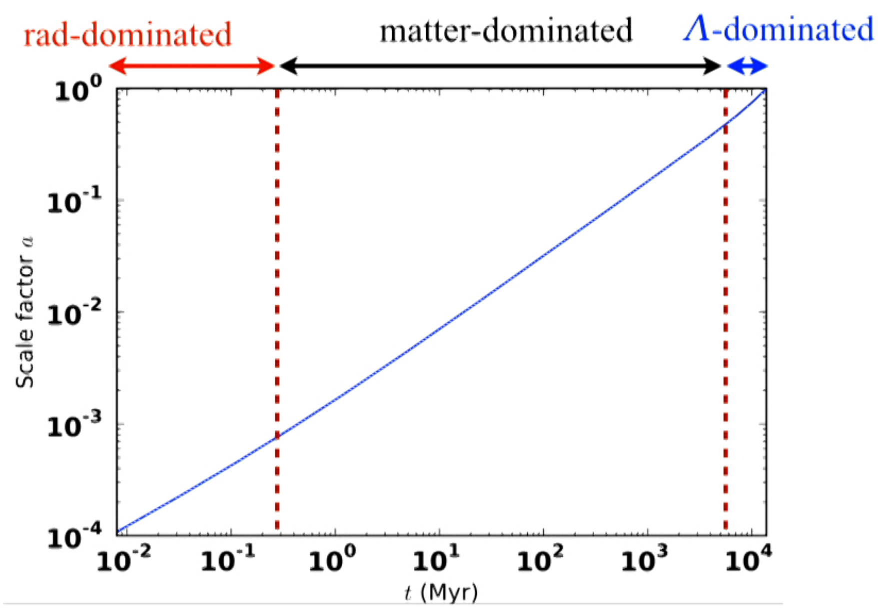
$\rho_{\rm de} = \text{constant}$	dark energy (assuming it is a c.c.!)
$\rho_{\rm m} \propto a(t)^{-3}$	nonrelativistic matter (dark + baryonic)
$\rho_{\rm rad} \propto a(t)^{-4}$	relativistic particles (photons, neutrinos)



"Expansion history is ΛCDM"

Scale factor a(t) evolves via Friedmann equation

$$\frac{d\log a}{dt} = \left(\frac{8\pi G}{3}\rho_{\rm tot}\right)^{1/2} = \left(\frac{8\pi G}{3}\left(\rho_{\rm de} + \rho_{\rm m}(a) + \rho_{\rm rad}(a)\right)\right)^{1/2}$$



The expansion history is parameterized by the first three parameters in the standard model (ρ_Λ, Ω_b, Ω_c).

$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}$ GeV^4	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

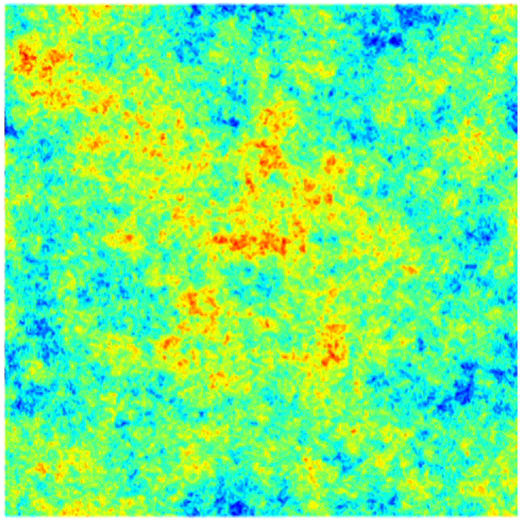
So far, we have not talked about perturbations. The next two parameters (Δ_ζ^2, n_s) specify the *initial* perturbations.

Initial conditions: at early times, the FRW metric has small perturbations.

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)} dx^2$$

The field $\zeta(x)$ is called the "adiabatic curvature" or the "initial curvature". This is a random field whose statistics can be described informally by the following statements:

- Initial perturbations are self-similar (no preferred scale)
- Almost scale-invariant, small trend toward more power on large scales.
- Characteristic size of fluctuations is $\Delta_\zeta \sim (5 \times 10^{-5})$



Initial conditions: at early times, the FRW metric has small perturbations.

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)} dx^2$$

More formally, $\zeta(x)$ is a Gaussian random field with the following power spectrum $P_\zeta(k)$. (This will be defined precisely later!)

$$\frac{k^3}{2\pi^2} P_\zeta(k) = \Delta_\zeta^2 \left(\frac{k}{0.05\ h\ \mathrm{Mpc}^{-1}} \right)^{n_s - 1}$$

with free parameters

$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index

"Initial perturbations are scalar adiabatic".

- "Scalar" means that there are no gravity wave perturbations in the initial metric. (Some models of inflation predict this, but so far it has not been observed.)

$$ds^2 = -dt^2 + a(t)^2 e^{2\zeta(x)}(\delta_{ij} + h_{ij}(x))$$

absent

- "Adiabatic" is more technical. It means that the ζ field also completely determines the perturbations in the stress-energy tensor, by a universal set of rules which will be explained later!

$$\rho(\mathbf{x}, t) = \bar{\rho}(t)\left(1 + \frac{4}{7}\zeta(\mathbf{x})\right)$$
$$\ldots$$

The statistics of the initial perturbations are parameterized by parameters (Δ_ζ^2, n_s) below.

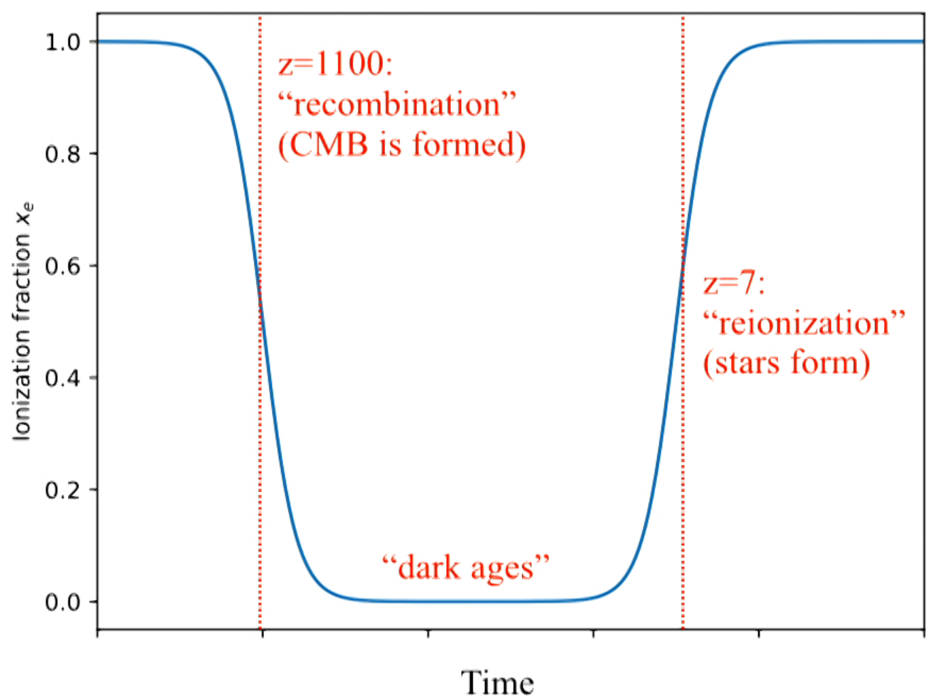
$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}$ GeV^4	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

The final parameter τ is an astrophysical nuisance parameter which we define for completeness.

Ionization history of the universe

$x_e(t)$ = electron ionization fraction

= probability that a random electron in the universe is ionized (rather than being part of an atom)



Ionization history of the universe

τ = CMB optical depth

= probability that a CMB photon emitted at $z \sim 1100$ scatters from an electron at low redshift, before being observed at $z=0$.

Astrophysical nuisance parameter: τ affects the CMB power spectrum.

When fitting cosmological parameters from the CMB, we need to include τ in the fit, and account for uncertainty in τ when assigning errors to other parameters.

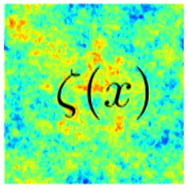
Standard model of cosmology:

- Background metric is FRW
- Expansion history is ΛCDM
- Initial perturbations are Gaussian random
- Initial perturbations are scalar adiabatic
- Power spectrum of initial perturbations is a power law: $(k^3/2\pi^2)P(k) = \Delta_\zeta^2 (k/k_0)^{n_s - 1}$

Six parameters:

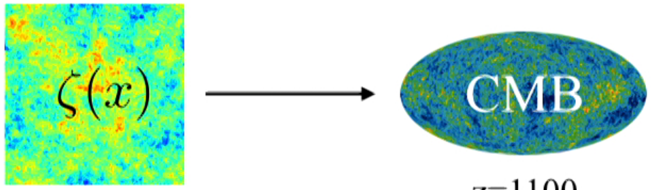
$\rho_\Lambda = (2.56 \pm 0.04) \times 10^{-47}\ \mathrm{GeV}^4$	Dark energy density (c.c.)
$\Omega_b = 0.0486 \pm 0.0007$	Baryonic(*) matter abundance
$\Omega_c = 0.267 \pm 0.009$	Cold dark matter abundance
$\Delta_\zeta^2 = (2.11 \pm 0.05) \times 10^{-9}$	Initial power spectrum amplitude
$n_s = 0.967 \pm 0.004$	Spectral index
$\tau = 0.058 \pm 0.012$	CMB optical depth

(*) "Baryons" = protons + neutrons + electrons(!)



High-z

The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field ζ(x) with a featureless power spectrum.

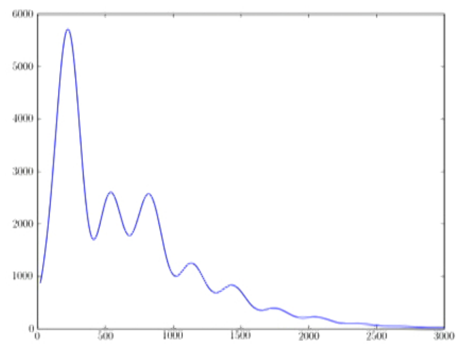


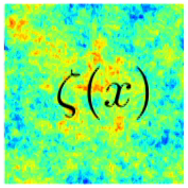
High-z

z=1100

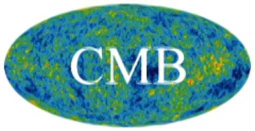
The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field ζ(x) with a featureless power spectrum.

As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.

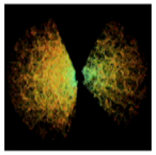




High-z



z=1100



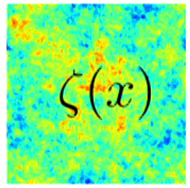
z~1

The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field ζ(x) with a featureless power spectrum.

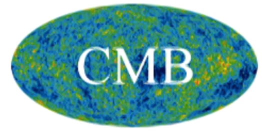
As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.

At late times (z~1), nonlinear effects are important and the perturbations are very non-Gaussian.

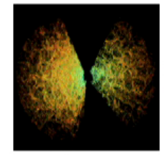
inflation?
cyclic universe?
something else?



High-z



z=1100



z~1

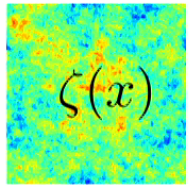
The standard cosmological model specifies the perturbations at very early times (high-z). They are fairly simple, and parameterized by a Gaussian random field ζ(x) with a featureless power spectrum.

As time evolves, the perturbations become more complex. By the time the CMB is formed (z=1100), a lot of physics has been "imprinted" on the power spectrum.

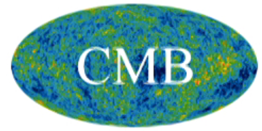
At late times (z~1), nonlinear effects are important and the perturbations are very non-Gaussian.

There are also models for the "early universe", a hypothetical phase preceding the radiation-dominated part of the expansion, which try to explain where the Gaussian field ζ came from.

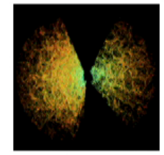
inflation?
cyclic universe?
something else?



High-z



z=1100



z~1

In each of these three stages, different physics is important:

- Early universe: Quantum mechanics in expanding spacetime generates Gaussian perturbations from vacuum
- Formation of the CMB: Linear perturbation theory in a plasma with multiple components (dark matter, baryons, photons, neutrinos) + metric degrees of freedom
- Late times: Gravitational N-body physics. Messy astrophysics! (galaxy formation, star formation, …)

Fundamental physics →

Cosmological constant ρ_Λ
Baryon abundance Ω_b
Dark matter abundance Ω_c
Initial amplitude Δ_ζ^2
Spectral index n_s
CMB optical depth τ

→ Data analysis

Challenge for observers: which model fits the data?

- ~1930: Expanding universe
- 1965: Big bang (discovery of CMB)
- ~1970: Dark matter
- 1992: Gaussian, nearly scale-invariant perturbations (COBE)
- 1998: Cosmological constant
- 2006: Deviation from scale invariance ($n_s < 1$)

Fundamental physics →

Cosmological constant ρ_Λ
Baryon abundance Ω_b
Dark matter abundance Ω_c
Initial amplitude Δ_ζ^2
Spectral index n_s
CMB optical depth τ

→ Data analysis

Fundamental physics →

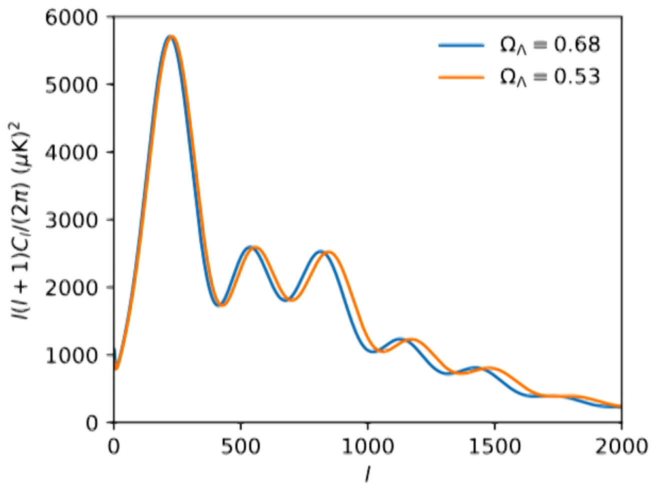
Cosmological constant ρ_Λ
Baryon abundance Ω_b
Dark matter abundance Ω_c
Initial amplitude Δ_ζ^2
Spectral index n_s
CMB optical depth τ

→ Data analysis

Challenge for theorists: explain this model at a fundamental level

- What is dark matter?
- Why is the cosmological constant so fine-tuned? (if late-time accelerated expansion is indeed a c.c.!)
- What physics is responsible for generating the initial Gaussian, nearly scale invariant fluctuations?

Cosmological observables (such as the CMB power spectrum) are sensitive to cosmological parameters, and can jointly constrain multiple parameters.



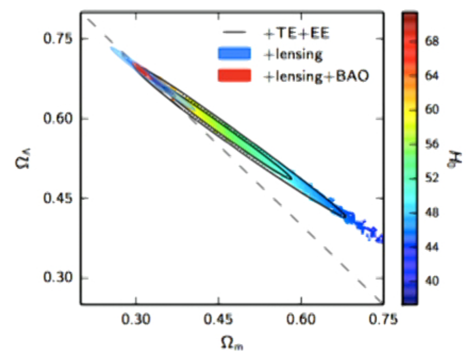


Fig. 26. Constraints in the Ω_m–Ω_Λ plane from the *Planck* TT+lowP data (samples; colour-coded by the value of H_0) and *Planck* TT,TE,EE+lowP (solid contours). The geometric degeneracy between Ω_m and Ω_Λ is partially broken because of the effect of lensing on the temperature and polarization power spectra. These limits are improved significantly by the inclusion of the *Planck* lensing reconstruction (blue contours) and BAO (solid red contours). The red contours tightly constrain the geometry of our Universe to be nearly flat.

Planck 2015

Cosmology is largely concerned with looking for extensions of the 6-parameter standard model.

- Non-Gaussian initial conditions
- Non-minimal neutrino mass
- Extra neutrino species or other light relics
- Interacting dark matter
- Nonzero spatial curvature
- Cosmological gravity waves

+ many others!

The standard model includes ingredients which were originally surprises (dark matter, cosmological constant, quantum mechanically generated perturbations).

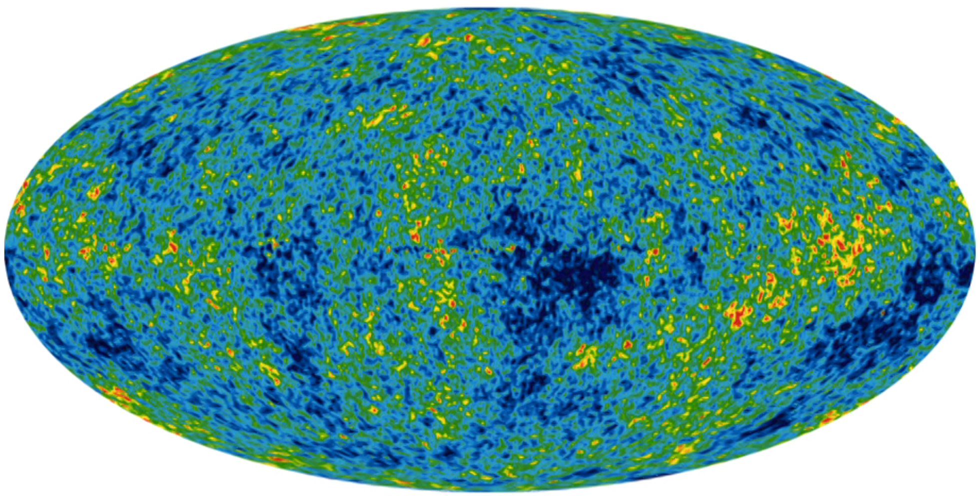
Will we find new surprises?

Part 2: random variables and fields

The standard model of cosmology is a probabilistic model.

For example, it can predict the probability of a given CMB realization occurring, but not the specific realization.

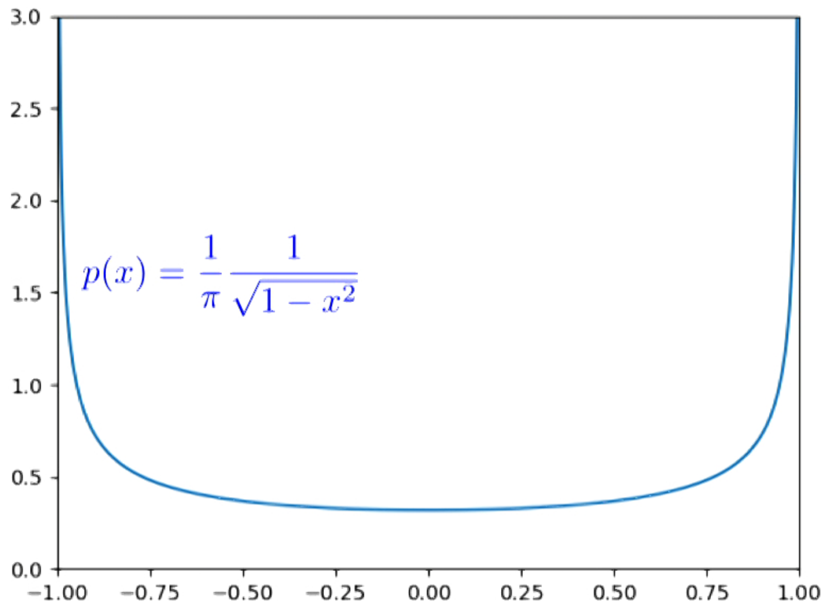
In this part of the lectures, we'll build up some machinery for working with random variables and fields.



Physicist's definition of a one-dimensional random variable X: anything with a probability distribution function (PDF) p(x).

The meaning of p(x) is "probability per unit x".

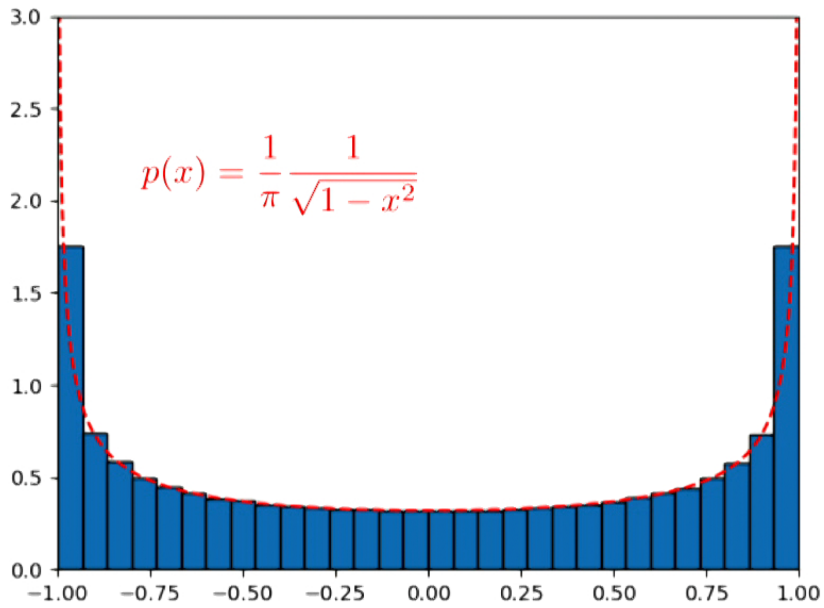
Here is an arbitrarily chosen example.

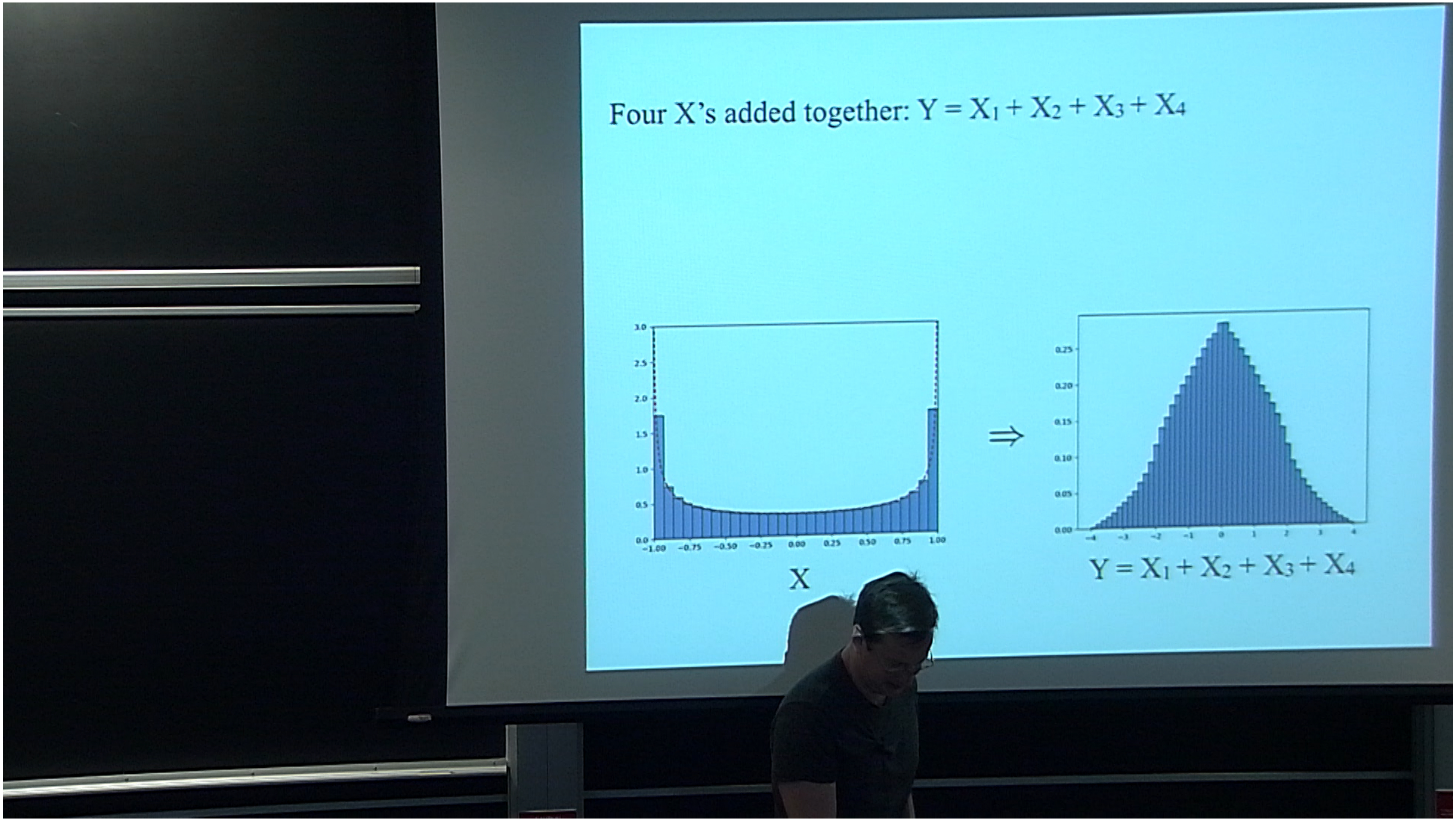


Histogram of 10^6 random samples in 30 bins, compared to the continuous PDF. The probability for the random variable X to be in bin [a,b] is:

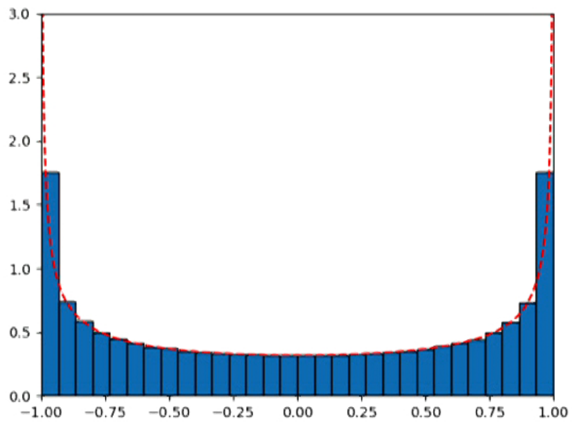
$$\mathrm{Prob}\Big(a < X < b\Big) = \int_a^b dx\, p(x)$$

Note that the PDF must satisfy $\int_{-\infty}^{\infty} dx\, p(x) = 1$



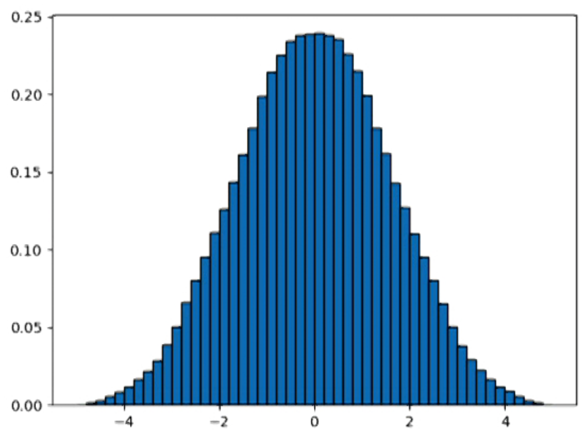

Four X's added together: Y = X1 + X2 + X3 + X4
X
⇒
Y = X1 + X2 + X3 + X4

Five X's added together: $Y = X_1 + X_2 + X_3 + X_4 + X_5$



X

$\Rightarrow$

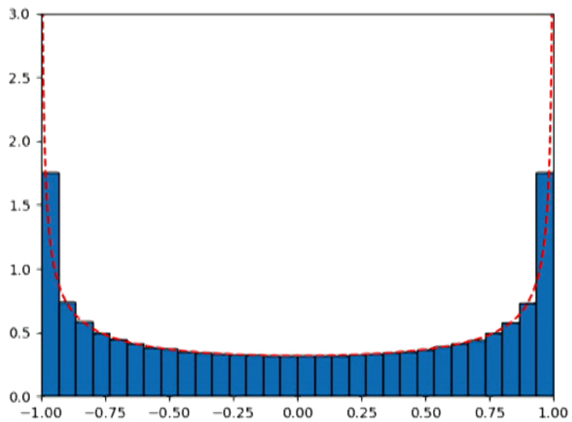


$Y = X_1 + X_2 + X_3 + X_4 + X_5$

Twenty X's added together: $Y = \sum_{i=1}^{20} X_i$

In the next few slides, we'll explain where the limiting PDF $p(x) = \frac{1}{\sqrt{20\pi}} e^{-x^2/20}$ comes from (including factors of 20, π).

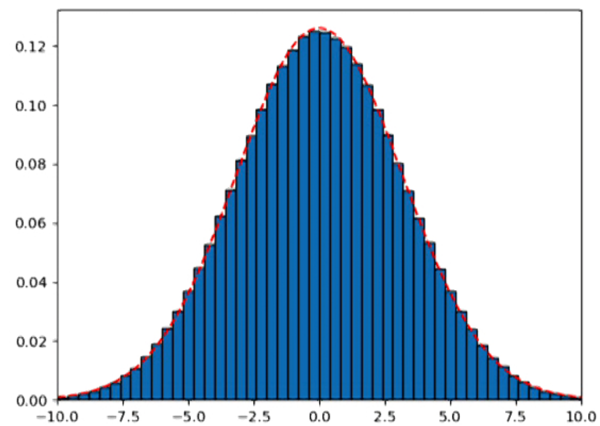
$$p(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}$$



X

$\Rightarrow$

$$p(x) = \frac{1}{\sqrt{20\pi}} e^{-x^2/20}$$



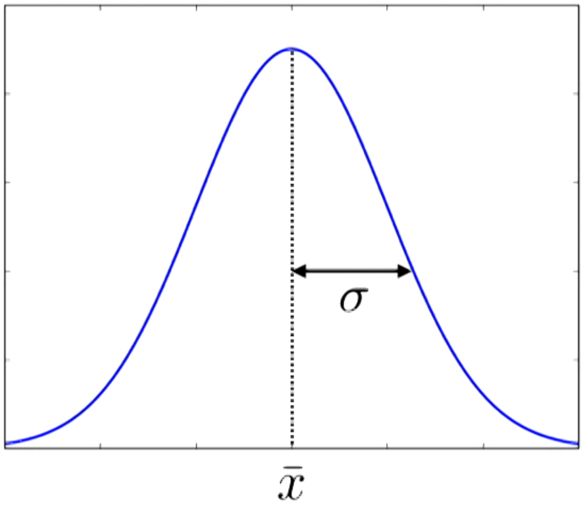
$$Y = \sum_{i=1}^{20} X_i$$

Central limit theorem: the sum of a large number of independent, identically distributed random variables has a PDF which is approximately Gaussian. (Proof omitted!)

The Gaussian PDF is defined by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

and has two parameters: a mean $\bar{x}$ and a width σ.



Some definitions: the mean and variance of a random variable X are defined by:

$$\bar{X} = \langle X \rangle \qquad \text{[mean]}$$

$$\begin{aligned} \mathrm{Var}(X) &= \langle X^2 \rangle - \langle X \rangle^2 \\ &= \langle (X - \bar{X})^2 \rangle \end{aligned} \qquad \text{[variance]}$$

$\sqrt{\mathrm{Var}(X)}$ can be interpreted as the "typical" size of fluctuations around the mean.

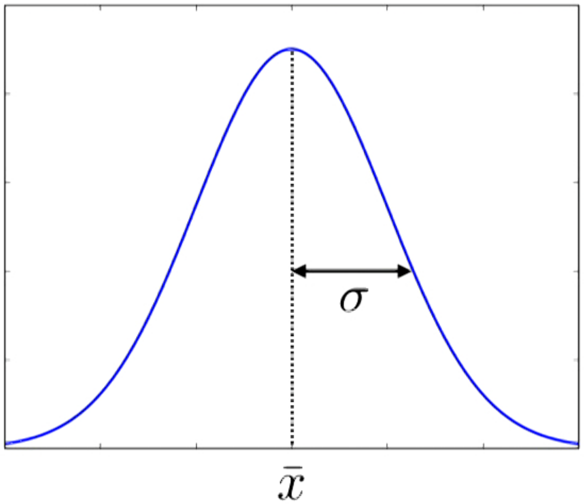
Example: For the Gaussian

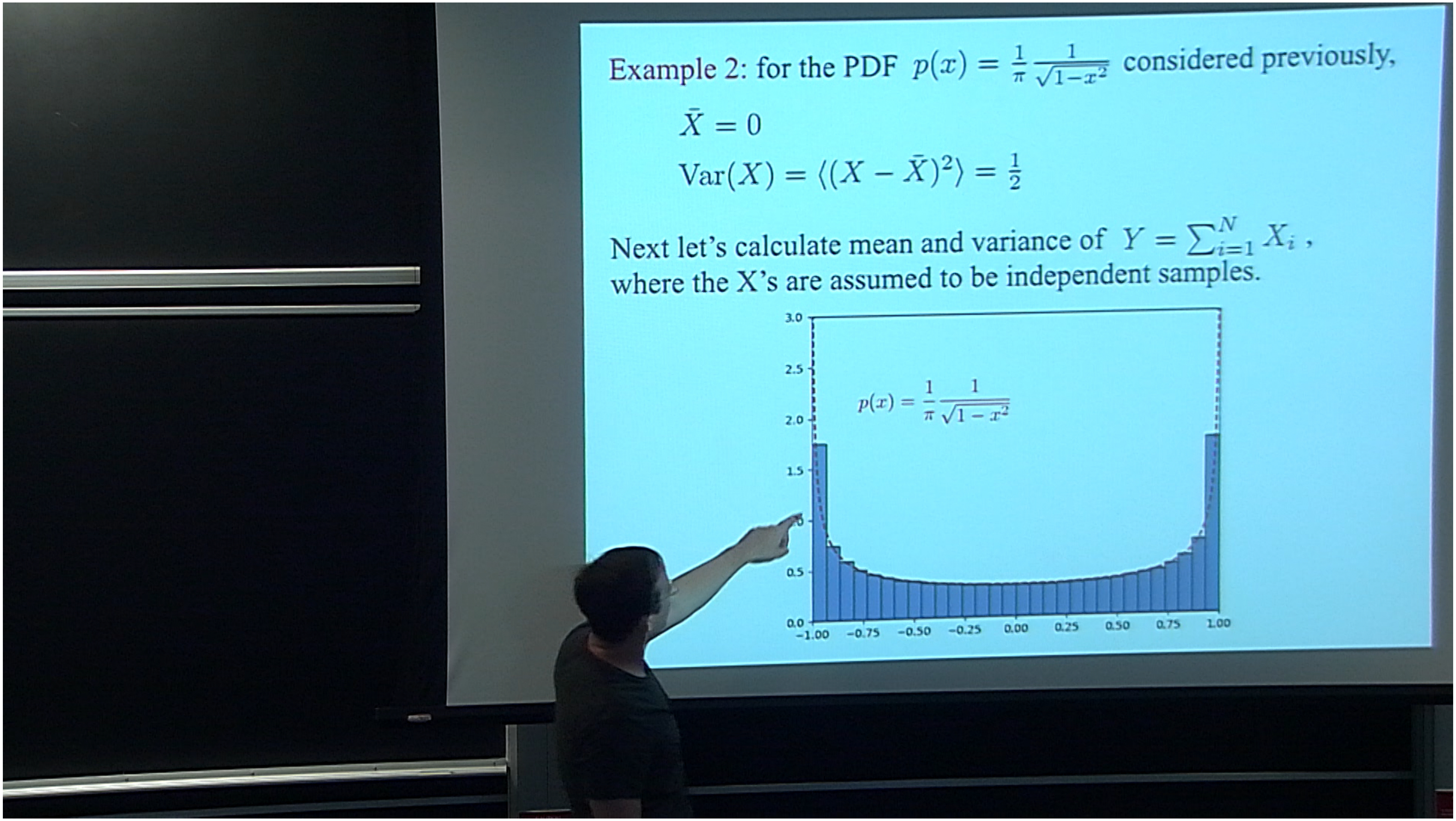
$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

a short calculation shows:

$$\text{Mean} = \int_{-\infty}^{\infty} dx\, p(x)\, x = \bar{x}$$

$$\text{Variance} = \int_{-\infty}^{\infty} dx\, p(x)\, (x^2 - \bar{x}^2) = \sigma^2$$




Example 2: for the PDF $p(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}$ considered previously,
$\bar{X} = 0$
$\mathrm{Var}(X) = \langle (X - \bar{X})^2 \rangle = \frac{1}{2}$
Next let's calculate mean and variance of $Y = \sum_{i=1}^{N} X_i$,
where the X's are assumed to be independent samples.
$p(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}$
3.0
2.5
2.0
1.5
0.5
0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00

Properties of expectation values:

$$\langle X \pm X' \rangle = \langle X \rangle \pm \langle X' \rangle$$

$$\langle cX \rangle = c\langle X \rangle$$ if c is a constant (not a random variable)

$$\langle X X' \rangle = \langle X \rangle \langle X' \rangle$$ if X, X' are independent random variables (not true in general!)

Now we can calculate mean and variance of $Y = \sum_{i=1}^{N} X_i$

$$\bar{Y} = \sum_{i=1}^{N} \bar{X}_i = 0$$

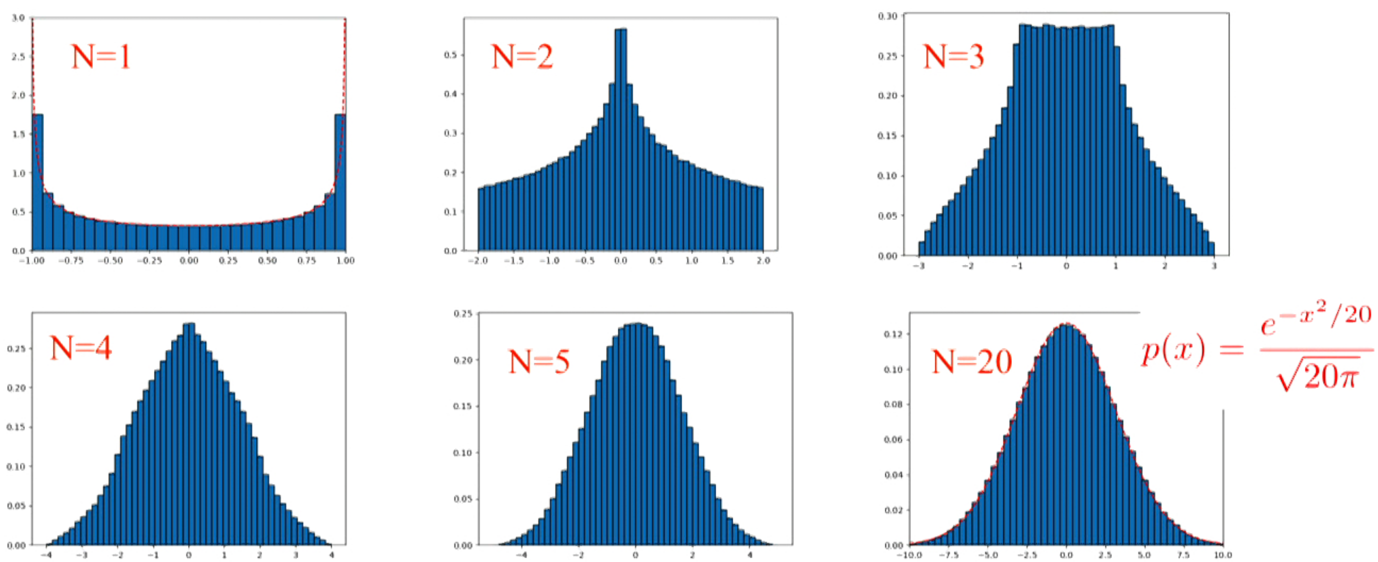
$$\begin{aligned} \mathrm{Var}(Y) &= \langle (Y - \bar{Y})^2 \rangle \\ &= \langle (\textstyle\sum_i X_i)^2 \rangle \\ &= \langle \textstyle\sum_i X_i^2 + \sum_{i \neq j} X_i X_j \rangle \\ &= \textstyle\sum_i \langle X_i^2 \rangle + \sum_{i \neq j} \langle X_i \rangle \langle X_j \rangle \\ &= N \left(\tfrac{1}{2} \right) \end{aligned}$$

This calculation gives the mean and variance of $Y = \sum_{i=1}^{N} X_i$:

$$\bar{Y} = 0 \qquad \mathrm{Var}(Y) = N/2 \qquad\qquad \text{(for all N)}$$

In general, the mean and variance do not determine the PDF p(x). However, for a Gaussian PDF they do!

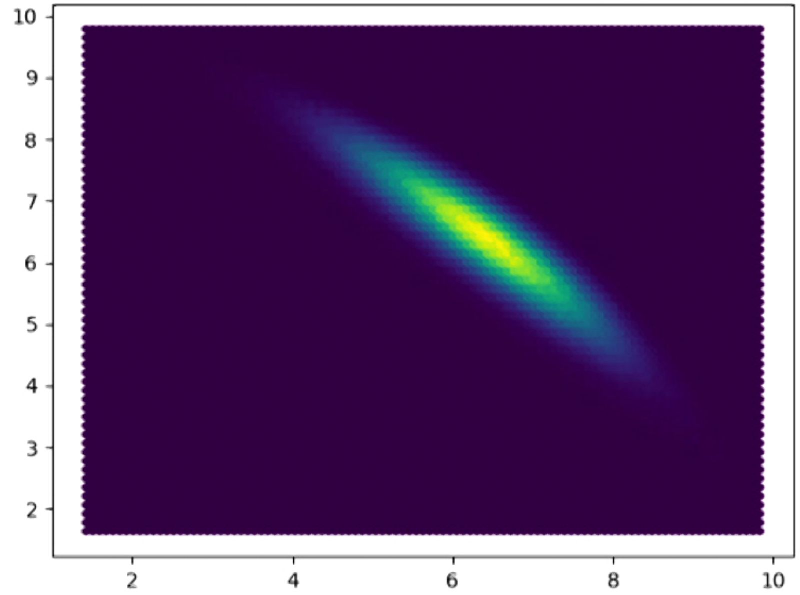
$$p(x) \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\bar{x})^2/2\sigma^2} = \frac{1}{\sqrt{\pi N}} e^{-x^2/N} \qquad \text{(for N >> 1)}$$



Multivariate random variables: let's generalize to the case of N random variables (X_1, …, X_N) which are not assumed independent.

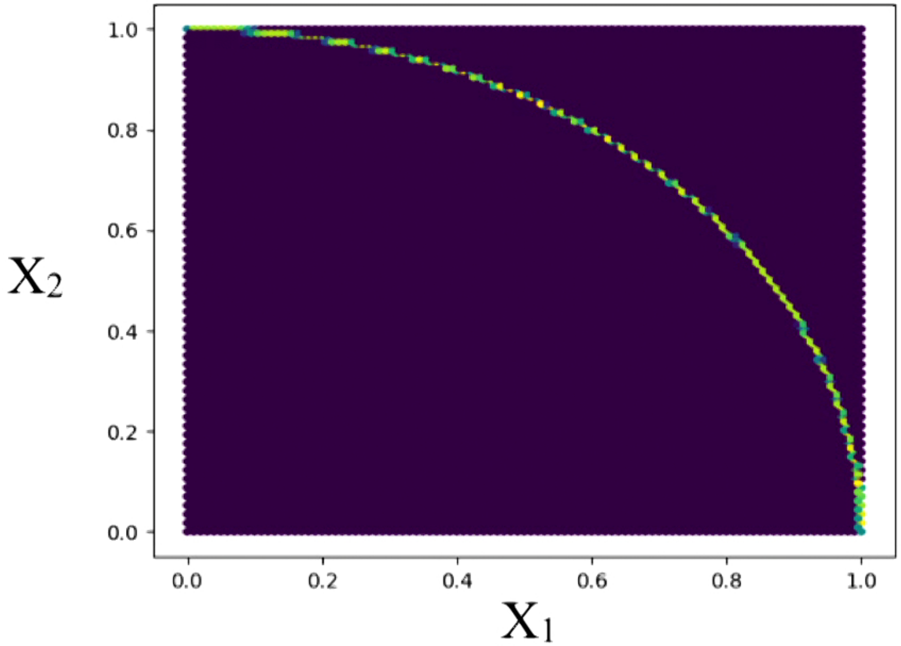
The PDF becomes a function of N variables $p(x_1,\ldots,x_N)$, and represents "probability per unit N-volume".

Example: a multivariate Gaussian (X_1, X_2) with a correlation between X_1 and X_2. (To be defined precisely in a few slides!)



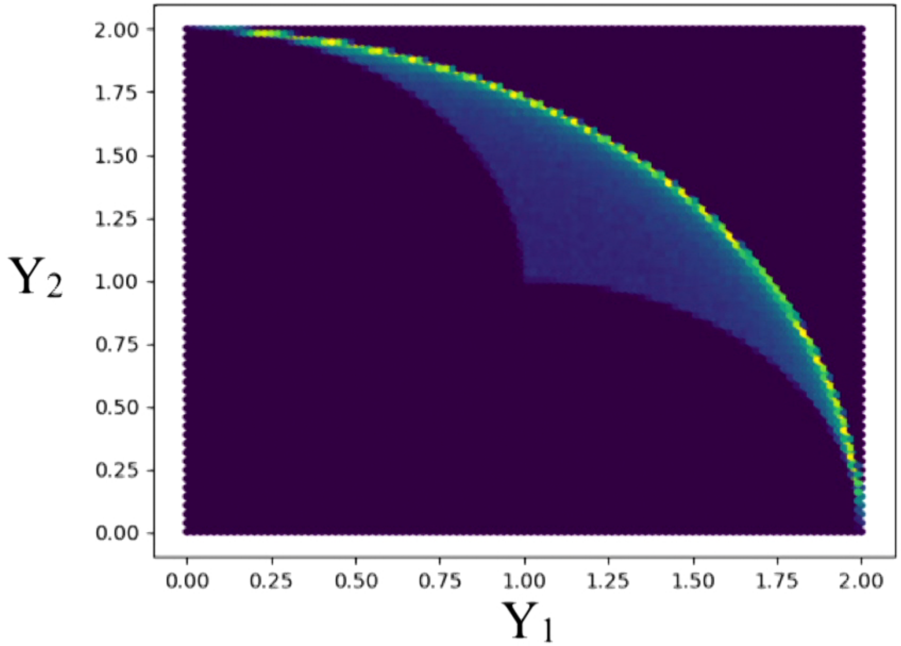
Example: $p(x_1, x_2) = \begin{cases} \frac{2}{\pi}\delta(\sqrt{x_1^2 + x_2^2} - 1) & \text{if } x_1, x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases}$

Just to show an extreme case where the variables x_1, x_2 are very non-independent!



Does the central limit theorem still hold when the random variable is a vector X_i? (In this case, a two-component vector)

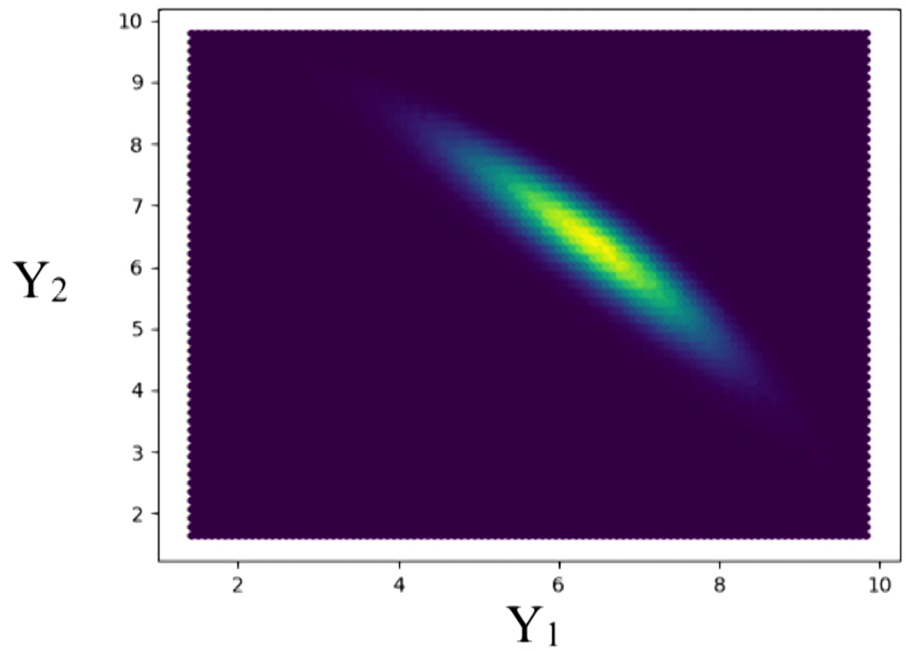
Two X's: $Y_i = X_i^{(1)} + X_i^{(2)}$



Ten X's: $Y_i = \sum_{j=1}^{10} X_i^{(j)}$

The distribution has become a multivariate Gaussian.

In two variables, the multivariate Gaussian has five parameters: two "means", and three parameters describing the size and orientation.



In N variables, the mean becomes an N-component vector

$$\bar{X}_i = \langle X_i \rangle$$

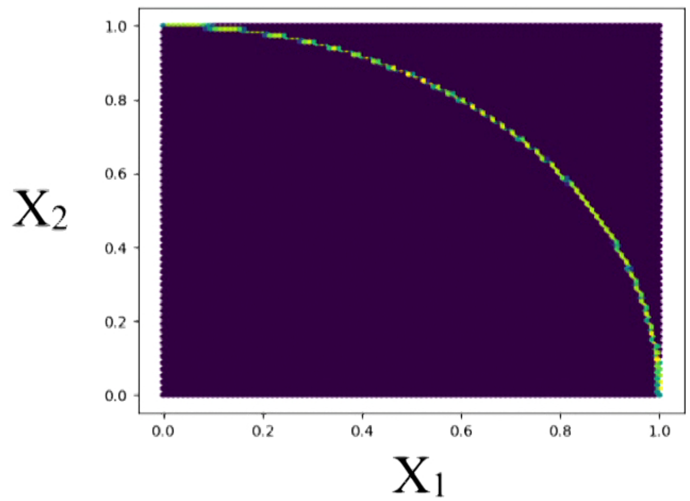
The variance generalizes to an N-by-N covariance matrix:

$$\mathrm{Cov}(X_i, X_j) = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle$$
$$= \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle$$

In our example, a short calculation gives the mean and covariance:

$$\begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \end{pmatrix} = \begin{pmatrix} 0.64 \\ 0.64 \end{pmatrix}$$

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix} = \begin{pmatrix} 0.095 & -0.087 \\ -0.087 & 0.095 \end{pmatrix}$$



Now we can give the definition of a multivariate Gaussian PDF:

$$p(x_1, \cdots, x_N) = \frac{1}{\mathrm{Det}(2\pi C)^{1/2}} \exp\left(-\frac{1}{2}(x_i - \bar{x}_i)C^{-1}_{ij}(x_j - \bar{x}_j)\right)$$

The PDF of a multivariate Gaussian random variable is determined by its mean $\bar{X}_i$ and covariance matrix $C_{ij} = \mathrm{Cov}(X_i, X_j)$

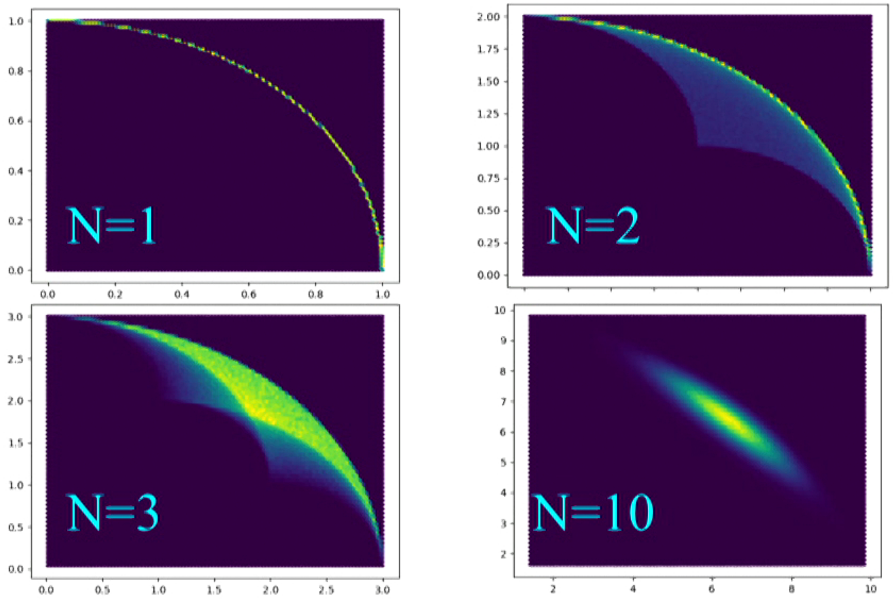
In cosmology, we are usually interested in Gaussian random variables. Therefore, it suffices to keep track of the mean (a vector) and the covariance (a matrix).

In this example:

$$\begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \end{pmatrix} = N \begin{pmatrix} 0.64 \\ 0.64 \end{pmatrix} \qquad \begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix} = N \begin{pmatrix} 0.095 & -0.087 \\ -0.087 & 0.095 \end{pmatrix}$$

In the large-N limit, these determine the PDF (central limit theorem):

$$p(x_1, x_2) \approx \frac{1}{\mathrm{Det}(2\pi C)^{1/2}} \exp\left(-\frac{1}{2}(x_i - \bar{x}_i) C^{-1}_{ij} (x_j - \bar{x}_j)\right) \qquad (N >> 1)$$



$$C_{ij} = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle$$
$$= \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle$$

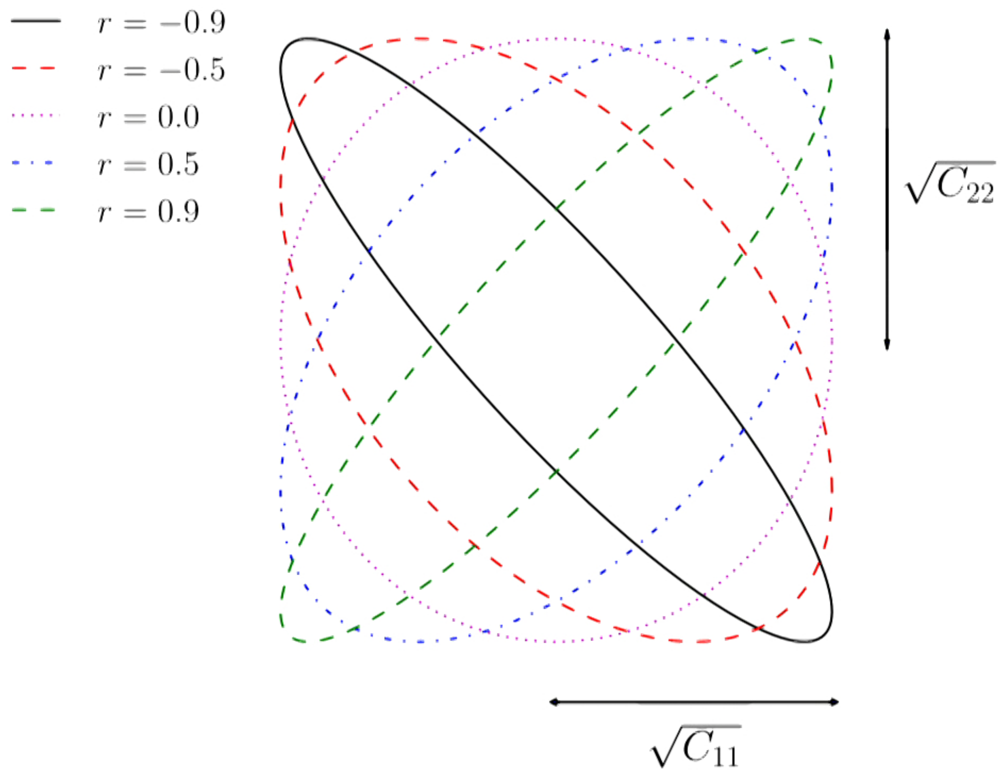
Diagonal elements C_{ii} of the covariance matrix are variances.
$C_{ii}^{1/2} \sim$ characteristic size of fluctuations in X_i around its mean.

Off-diagonals C_{ij} quantify the level of correlation between random variables X_i, X_j. The correlation coefficient

$$\mathrm{Corr}(X_i, X_j) = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}$$

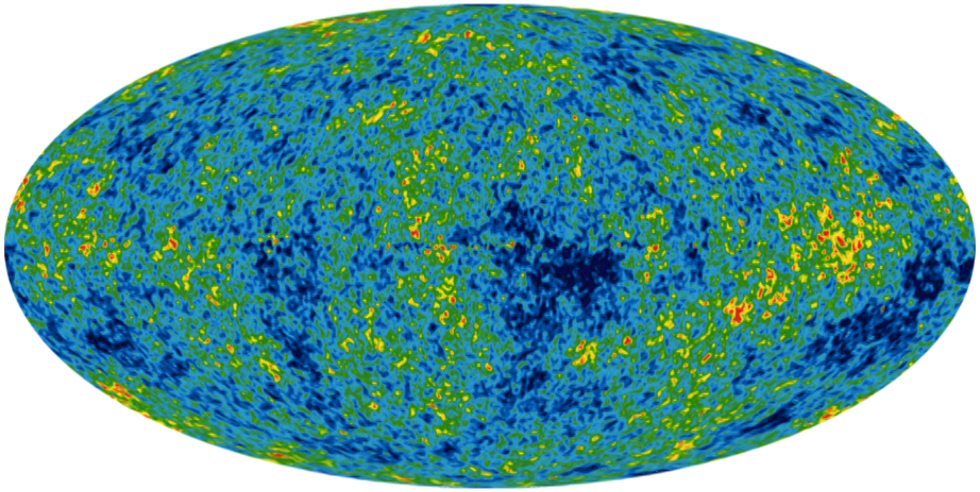
is always between -1 and 1.

Visual representation of covariance matrix (where $r = \dfrac{C_{12}}{\sqrt{C_{11}C_{22}}}$)



The CMB is a multivariate Gaussian random variable!

If the map below is represented with $N=10^7$ pixels, then the statistics are described perfectly (as far as we know) by a multivariate Gaussian, whose N-by-N covariance matrix can be calculated numerically in the standard model.



Behavior of mean and covariance under linear transformations.

Let X_i be an N-component random variable, and define an M-component random variable Y_a by:

$$Y_a = A_{ai} X_i \qquad (A_{ai} \text{ is an M-by-N matrix})$$

Then the mean (a vector) and covariance matrix transform as:

$$\bar{Y}_a = \langle A_{ai} X_i \rangle = A_{ai} \bar{X}_i$$

$$\begin{aligned} \mathrm{Cov}(Y_a, Y_b) &= \langle (Y_a - \bar{Y}_a)(Y_b - \bar{Y}_b) \rangle \\ &= \langle (A_{ai}(X_i - \bar{X}_i))(A_{bj}(X_j - \bar{X}_j)) \rangle \\ &= A_{ai} A_{bj} \langle (X_i - \bar{X}_i)(X_j - \bar{X}_j) \rangle \\ &= A_{ai} A_{bj} \mathrm{Cov}(X_i, X_j) \end{aligned}$$

Or in index-free notation:

$$\bar{Y} = A\bar{X} \qquad C_Y = A C_X A^T$$

For arbitrary random variables $Y_a = A_{ai} X_i$, the mean and covariance transform as:

$$\bar{Y} = A\bar{X} \qquad C_Y = A C_X A^T$$

Theorem (proof omitted): if X_i is Gaussian, then Y_a is also Gaussian. In this case, the mean and covariance completely determine the statistics.

In particular, the question of whether a random variable X_i is Gaussian does not depend on the choice of basis. (Changing basis $X_i \rightarrow X'_i$ is the special case where A is invertible.)

Sometimes, problems involving random variables are linear algebra problems in disguise.

Example: how to simulate (on the computer) a Gaussian random variable X_i with specified covariance matrix C_{ij}?

Sometimes, problems involving random variables are linear algebra problems in disguise.

Example: how to simulate (on the computer) a Gaussian random variable X_i with specified covariance matrix C_{ij}?

Answer: diagonalize C

$$C = R \Lambda R^{-1} \qquad \text{where } \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_N \end{pmatrix} \text{ and } R^{-1} = R^T$$

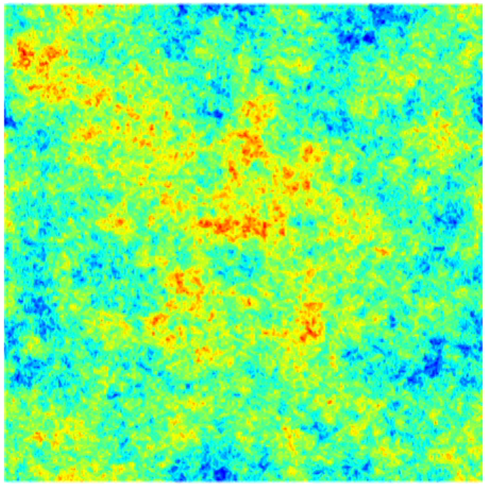
Now simulate a Gaussian random variable Y_i with covariance matrix Λ (straightforward, since Λ is diagonal).

Define $X = RY$. This is a Gaussian random variable with covariance matrix $C_X = R\, C_Y\, R^T = R\, \Lambda\, R^{-1} = C$, as desired.

Random fields.

Consider an image f_p with 256^2 (say) pixels. (where p=1, …, 256^2).

If f_p is a random variable, then its covariance $C_{pp'}$ is a 256^2-by-256^2 matrix. (Assume mean $\bar{f}_p = 0$ for simplicity.)



Now take the continuum limit:

pixelized image f_p	→	continuous function f(**x**)
covariance matrix $C_{pq} = \langle f_p f_{p'} \rangle$	→	two-point correlation function $\langle f(\mathbf{x})\, f(\mathbf{x}') \rangle$

Unless stated otherwise, we will be interested in random fields which are translation and rotation invariant, so that the two-point function < f(**x**) f(**x'**)> depends only on the *scalar* separation |**x** - **x'**|.

$$\langle f(\mathbf{x}) f(\mathbf{x}')\rangle = \zeta(|\mathbf{x} - \mathbf{x}'|)$$

ζ is called the "correlation function".

Now let's compute the two-point function in Fourier space.

$$\begin{aligned}
\langle f(\mathbf{k})f(\mathbf{k}')^*\rangle &= \left\langle \left(\int d^n\mathbf{x}\ f(x)e^{-i\mathbf{k}\cdot\mathbf{x}}\right)\left(\int d^n\mathbf{x}'\ f(x)e^{i\mathbf{k}'\cdot\mathbf{x}'}\right)\right\rangle \\
&= \int d^n\mathbf{x}\,d^n\mathbf{x}'\left\langle f(\mathbf{x})f(\mathbf{x}')\right\rangle e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'} \\
&= \int d^n\mathbf{x}\,d^n\mathbf{x}'\,\zeta(|\mathbf{x}-\mathbf{x}'|)e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot\mathbf{x}'} \\
&= \int d^n\mathbf{x}\,d^n\mathbf{r}\,\zeta(|\mathbf{r}|)e^{-i\mathbf{k}\cdot\mathbf{x}+i\mathbf{k}'\cdot(\mathbf{x}-\mathbf{r})} \qquad\qquad (\mathbf{r}=\mathbf{x}-\mathbf{x}') \\
&= \left[\int d^n\mathbf{r}\,\zeta(|\mathbf{r}|)e^{-i\mathbf{k}\cdot\mathbf{r}}\right](2\pi)^n\delta^n(\mathbf{k}-\mathbf{k}')
\end{aligned}$$

The quantity in brackets is called the power spectrum P(k).

We have now shown that the two-point statistics of a random field are given equivalently by:

$$\langle f(\mathbf{x}) f(\mathbf{x}')\rangle = \zeta(|\mathbf{x} - \mathbf{x}'|) \qquad \text{in real space}$$

$$\left\langle f(\mathbf{k}) f(\mathbf{k}')^* \right\rangle = P(|\mathbf{k}|)\,(2\pi)^n \delta^n(\mathbf{k} - \mathbf{k}') \qquad \text{in Fourier space}$$

and the correlation function ζ(r) and power spectrum P(k) are related to each other by Fourier transforms ("Weiner-Khinchin theorem"):

$$P(k) = \int d^n\mathbf{r}\,\zeta(|\mathbf{r}|)\,e^{-i\mathbf{k}\cdot\mathbf{r}}$$

$$\zeta(r) = \int \frac{d^n\mathbf{k}}{(2\pi)^n}\,P(|\mathbf{k}|)\,e^{i\mathbf{k}\cdot\mathbf{r}}$$

A random field is Gaussian if its real-space values f(**x**) are a multivariate Gaussian random variable in the usual sense. In this case, the statistics are completely determined by the two-point function (either $\zeta(r)$ or P(k)).

Gaussian random fields are easy to think about in Fourier space, since the covariance is always diagonal:

$$\left\langle f(\mathbf{k}) f(\mathbf{k}')^* \right\rangle = P(|\mathbf{k}|)\,(2\pi)^n \delta^n(\mathbf{k}-\mathbf{k}') \qquad \textbf{(*)}$$

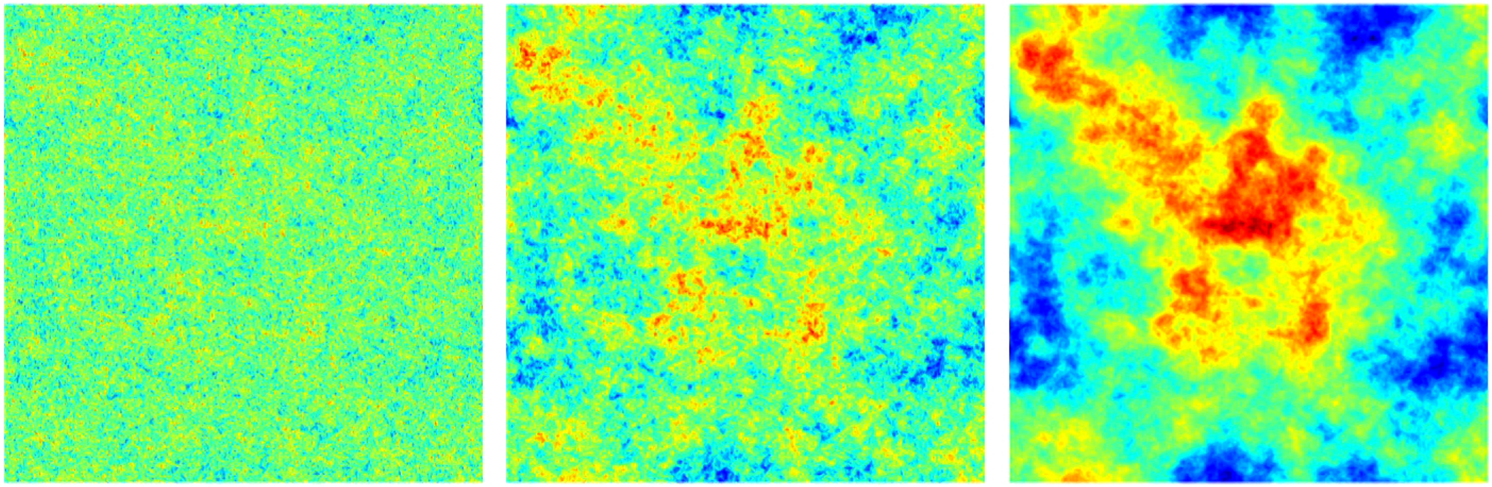
In Fourier space, a Gaussian random field is just a collection of *independent* Gaussian random variables f(**k**).

The delta function on the RHS of (*) can also be understood from translation invariance. Under a translation x => x + a, the two-point function transforms as:

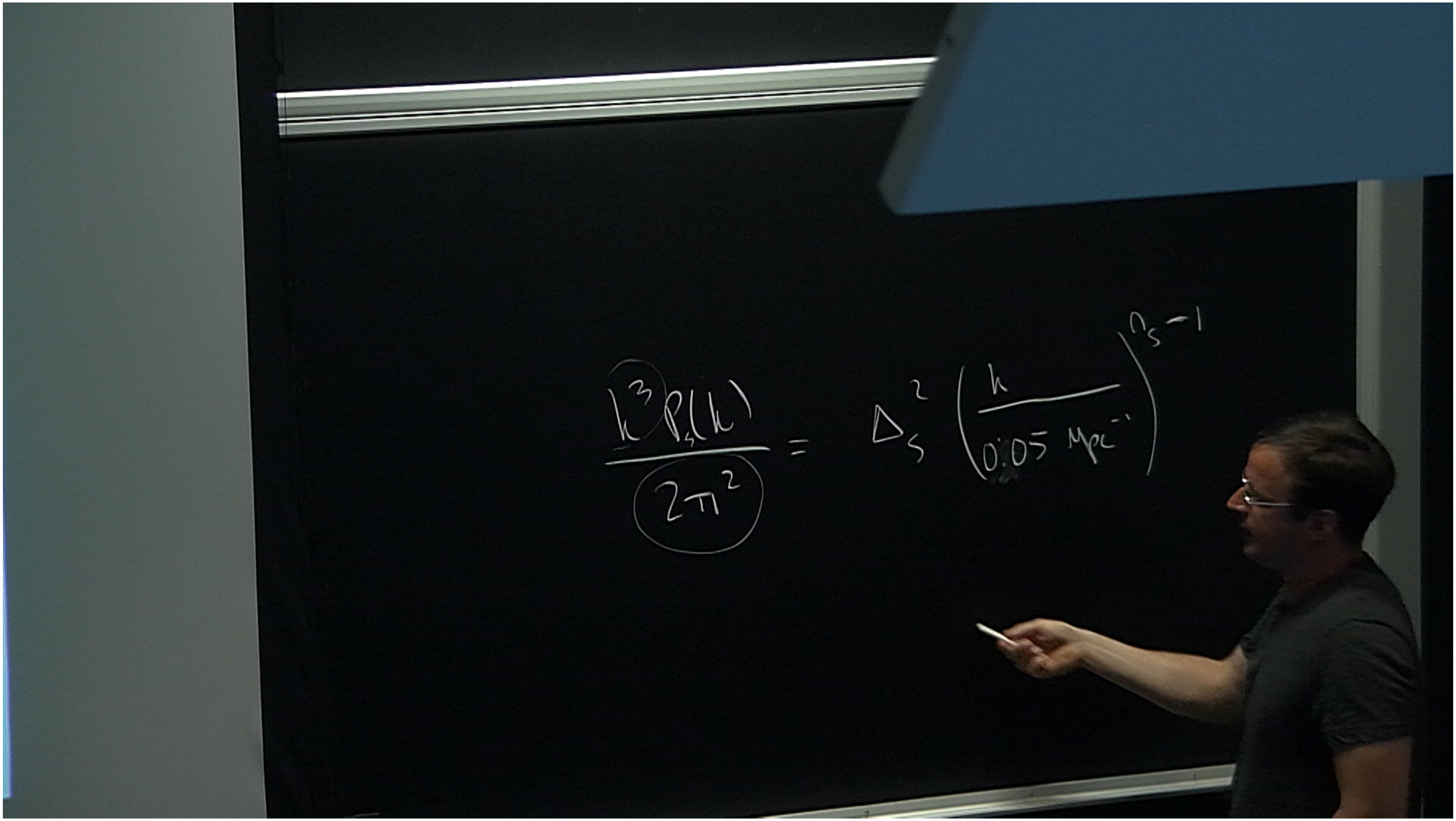
$$\left\langle f(\mathbf{k}) f(\mathbf{k})^* \right\rangle \to e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{a}} \left\langle f(\mathbf{k}) f(\mathbf{k})^* \right\rangle$$

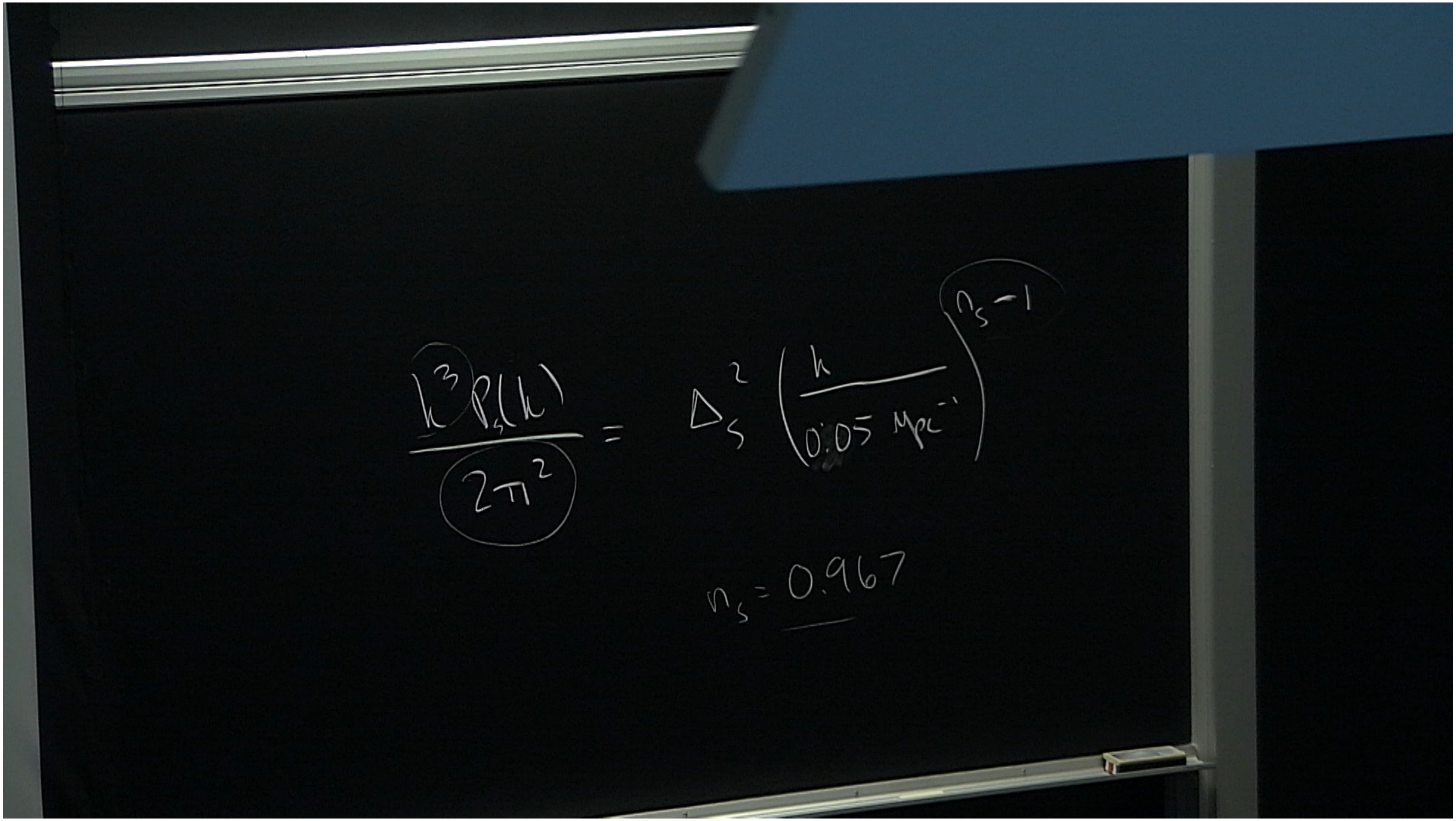
Example: Two-dimensional Gaussian random fields with power-law spectra $P(l) \propto l^{\alpha}$

(Note: cosmologists are hardwired to denote wavenumbers by k in 3D, by l in 2D, and by ω in 1D.)



$\alpha = -1$ "blue" spectrum | $\alpha = -2$ scale invariant | $\alpha = -3$ "red" spectrum

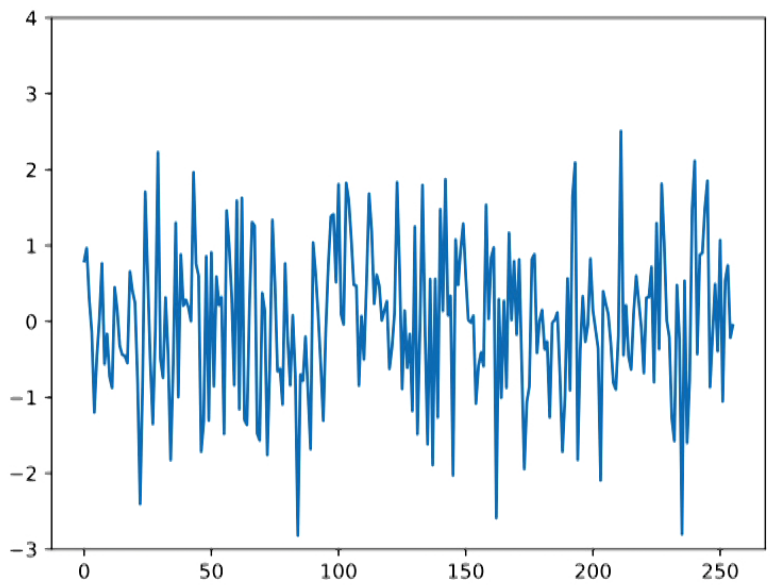

$\frac{k^3 P_s(k)}{2\pi^2} = \Delta_s^2 \left(\frac{k}{0.05\ \mathrm{Mpc}^{-1}}\right)^{n_s-1}$


$\frac{k^3 P_s(k)}{2\pi^2} = \Delta_s^2 \left(\frac{k}{0.05\ \mathrm{Mpc}^{-1}}\right)^{n_s - 1}$
$n_s = 0.967$

Gaussian white noise: simplest example of a Gaussian random field. The correlation function is a delta function.

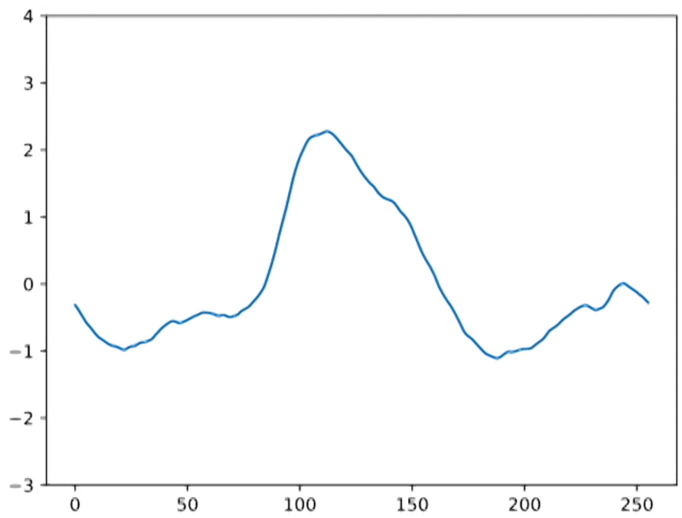
$$\zeta(\mathbf{r}) = A\,\delta^n(\mathbf{r})$$

Each pixel value is an independent Gaussian random variable. (Covariance matrix is diagonal in real space and Fourier space!)



A linear operator applied to a GRF (Gaussian random field) gives another GRF. (This follows from the general statement that linear combinations of Gaussians are Gaussian.)

Example: what is the power spectrum of a one-dimensional Gaussian random walk? (Obtained by adding an independent Gaussian random number at each timestep.)



A linear operator applied to a GRF (Gaussian random field) gives another GRF. (This follows from the general statement that linear combinations of Gaussians are Gaussian.)

Example: what is the power spectrum of a one-dimensional Gaussian random walk? (Obtained by adding an independent Gaussian random number at each timestep.)

To answer this, we note that a random walk is the integral of white noise. Therefore:

$$
\begin{aligned}
f_{\mathrm{RW}}(\omega) &= \frac{1}{i\omega}\, f_{\mathrm{WN}}(\omega) \\
P_{\mathrm{RW}}(\omega) &= \frac{1}{\omega^2}\, P_{\mathrm{WN}}(\omega) \\
&= \frac{A}{\omega^2}
\end{aligned}
$$

Another example which is more representative of the CMB.
Let f(t,x) be a field which evolves via the wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - c_s^2 \frac{\partial^2}{\partial x^2} \right) f(t,x) = 0$$

c_s = “sound speed”

with the following initial conditions at t=0:

- f(x) is a Gaussian random field with power spectrum $P_0(k)$
  $\partial f / \partial t = 0$

Question: what is the power spectrum $P_T(k)$ of a spatial “snapshot” f(T,x) at time t=T?

We take a spatial Fourier transform x → k (but not t→w).
Then the wave equation $(\partial_t^2 - c_s^2 \partial_x^2) f = 0$ becomes:

$$\left(\frac{\partial^2}{\partial t^2} + c_s^2 k^2 \right) f(t,k) = 0$$

and the solution is (using $\partial f / \partial t = 0$)

$$f(t,k) = \cos(c_s k t) f(0,k)$$

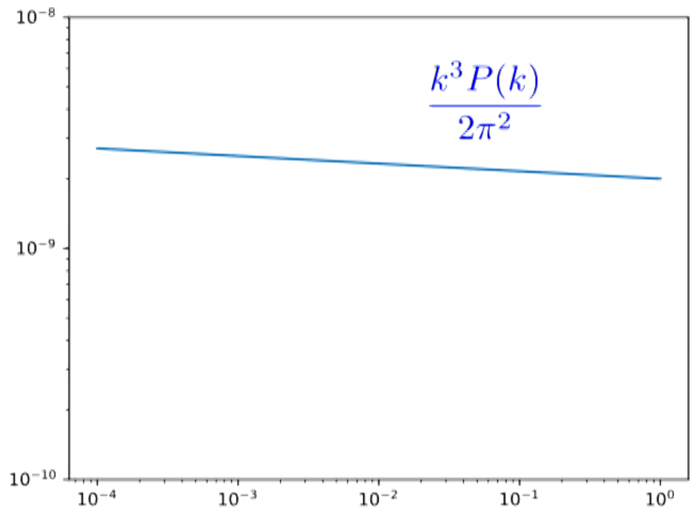
The spatial power spectrum $P_T(k)$ at time t=T is:

$$P_T(k) = \cos(c_s k T)^2 P_0(k)$$

i.e. time evolution imprints peaks on the power spectrum.

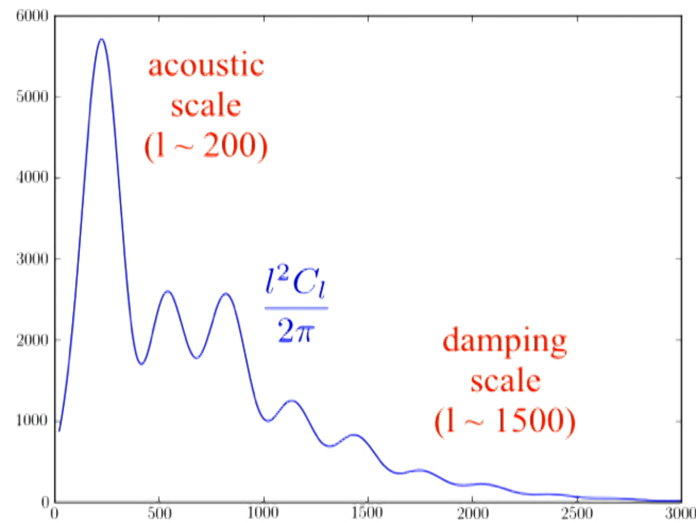
Analogously, time evolution "imprints" features on cosmological power spectra, starting from a featureless initial power spectrum.

3D power spectrum of initial conditions (adiabatic curvature)

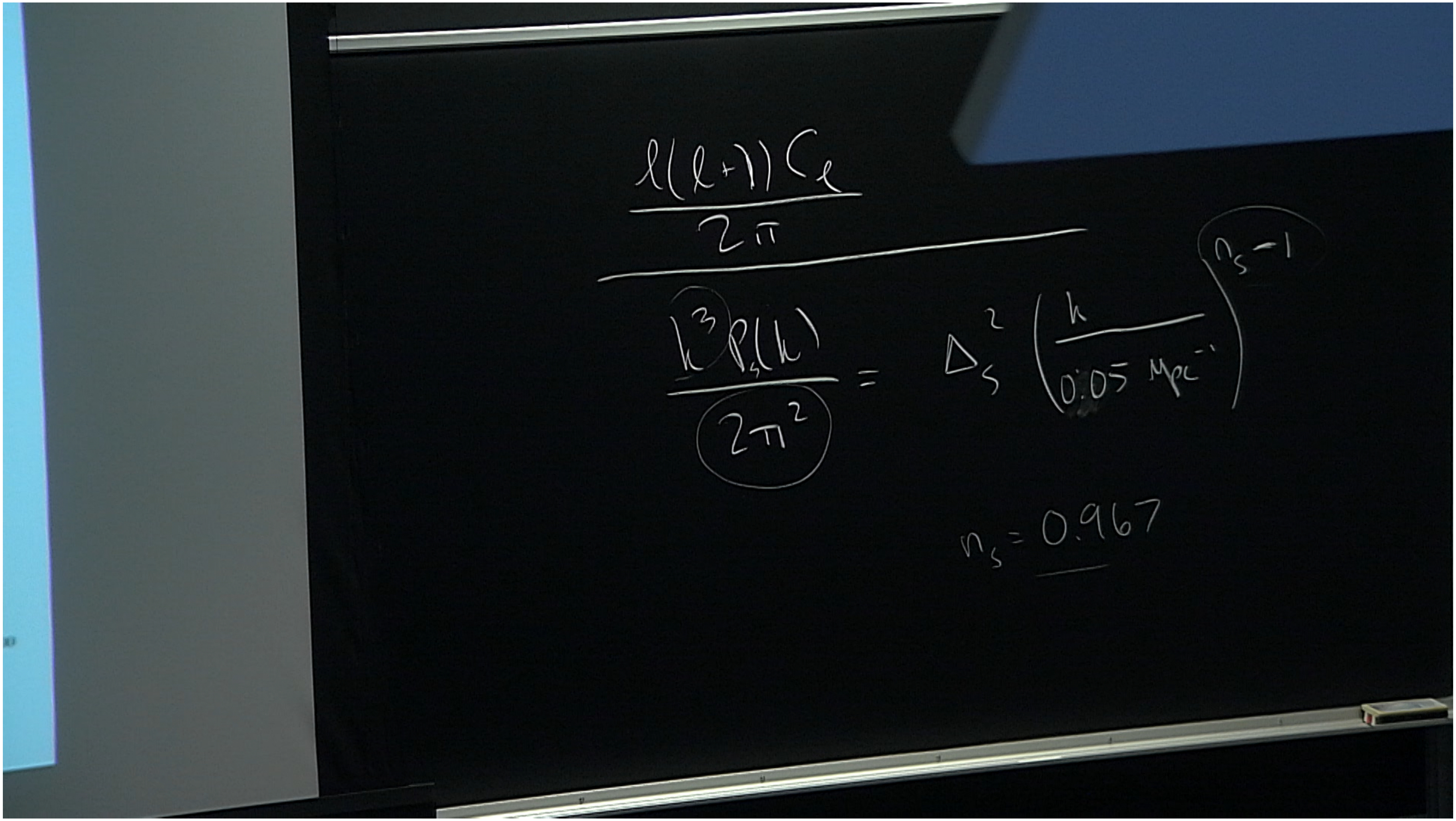


wavenumber k (Mpc^{-1})

2D CMB power spectrum



angular multipole l


l(l+1) C_l / 2π
k^3 P_s(k) / 2π^2 = Δ_s^2 (k / 0.05 Mpc^-1)^(n_s - 1)
n_s = 0.967

We take a spatial Fourier transform x → k (but not t→w).
Then the wave equation $(\partial_t^2 - c_s^2\partial_x^2)f = 0$ becomes:

$$\left(\frac{\partial^2}{\partial t^2} + c_s^2 k^2\right) f(t,k) = 0$$

and the solution is (using $\partial f/\partial t = 0$)

$$f(t,k) = \cos(c_s k t) f(0,k)$$

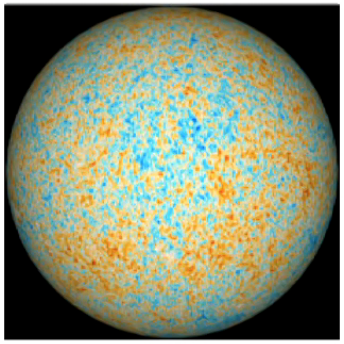
The spatial power spectrum $P_T(k)$ at time t=T is:

$$P_T(k) = \cos(c_s k T)^2 P_0(k)$$

i.e. time evolution imprints peaks on the power spectrum.

Curved sky.

So far, our fields have been defined on Euclidean space, but some fields are defined on the unit sphere, e.g. CMB temperature $T(\theta,\phi)$.



In Euclidean space, any field $f(x)$ can be represented as a linear combination of plane waves e^{ikx} (Fourier transform).

Analogous statement on the sphere: any field $f(\theta,\phi)$ is a linear combination of spherical harmonics $Y_{lm}(\theta,\phi)$.

The spherical harmonic $Y_{lm}(\theta,\phi)$ is a special function defined for integers $\ell = 0,1,2,\ldots$ and $m = -\ell, (-\ell+1), \ldots, \ell$.

Spherical analogue of a plane wave e^{ikx}. The wavenumber ℓ is quantized (an integer), and there are $(2\ell+1)$ harmonics for each ℓ.

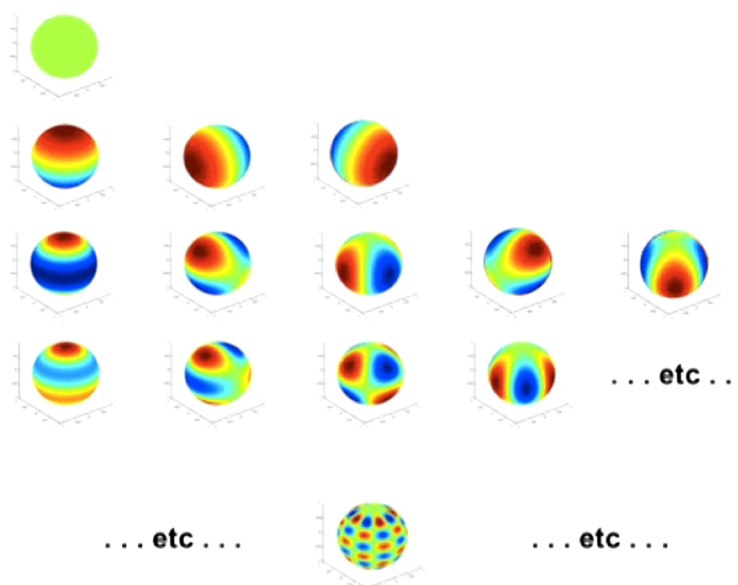
Any function $f(\theta,\phi)$ is representable as $f(\theta,\phi) = \sum_{lm} a_{lm} Y_{lm}(\theta,\phi)$

$\ell=0$ (monopole)

$\ell=1$ (dipole)

$\ell=2$ (quadrupole)

$\ell=3$ (octopole)



Euclidean field	Spherical field
real-space representation	
$f(\mathbf{x})$	$f(\theta, \phi)$
harmonic-space representation	
$\tilde{f}(\mathbf{k})$	a_{lm}
harmonic transform	
$f(\mathbf{x}) = \int \frac{d^n\mathbf{k}}{(2\pi)^n} \tilde{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}$	$f(\theta, \phi) = \sum_{lm} a_{lm} Y_{lm}(\theta, \phi)$
inverse transform	
$\tilde{f}(\mathbf{k}) = \int d^n\mathbf{x}\, f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}$	$a_{lm} = \int d(\cos\theta)\, d\phi\, f(\theta, \phi) Y^*_{lm}(\theta, \phi)$
power spectrum	
$\langle f(\mathbf{k}) f(\mathbf{k}')^* \rangle = P(k)\,(2\pi)^n \delta^n(\mathbf{k} - \mathbf{k}')$	$\langle a_{lm} a^*_{l'm'} \rangle = C_l \delta_{ll'} \delta_{mm'}$

Part 3: forecasting and the Fisher matrix