Astro/Phys 224 Spring 2014

Origin and Evolution of the Universe

Week 10

Cosmic Inflation: Before & After

Joel Primack

University of California, Santa Cruz

Physics 224 - Spring 2014 – Tentative Term Project Topics

Tuesday Afternoon June 10 - 1-4 pm - ISB 231

Adam Coogan – The BICEP2 Result and Its Implications for Inflationary Cosmology

Devon Hollowood – Some Aspect of the Dark Energy Survey

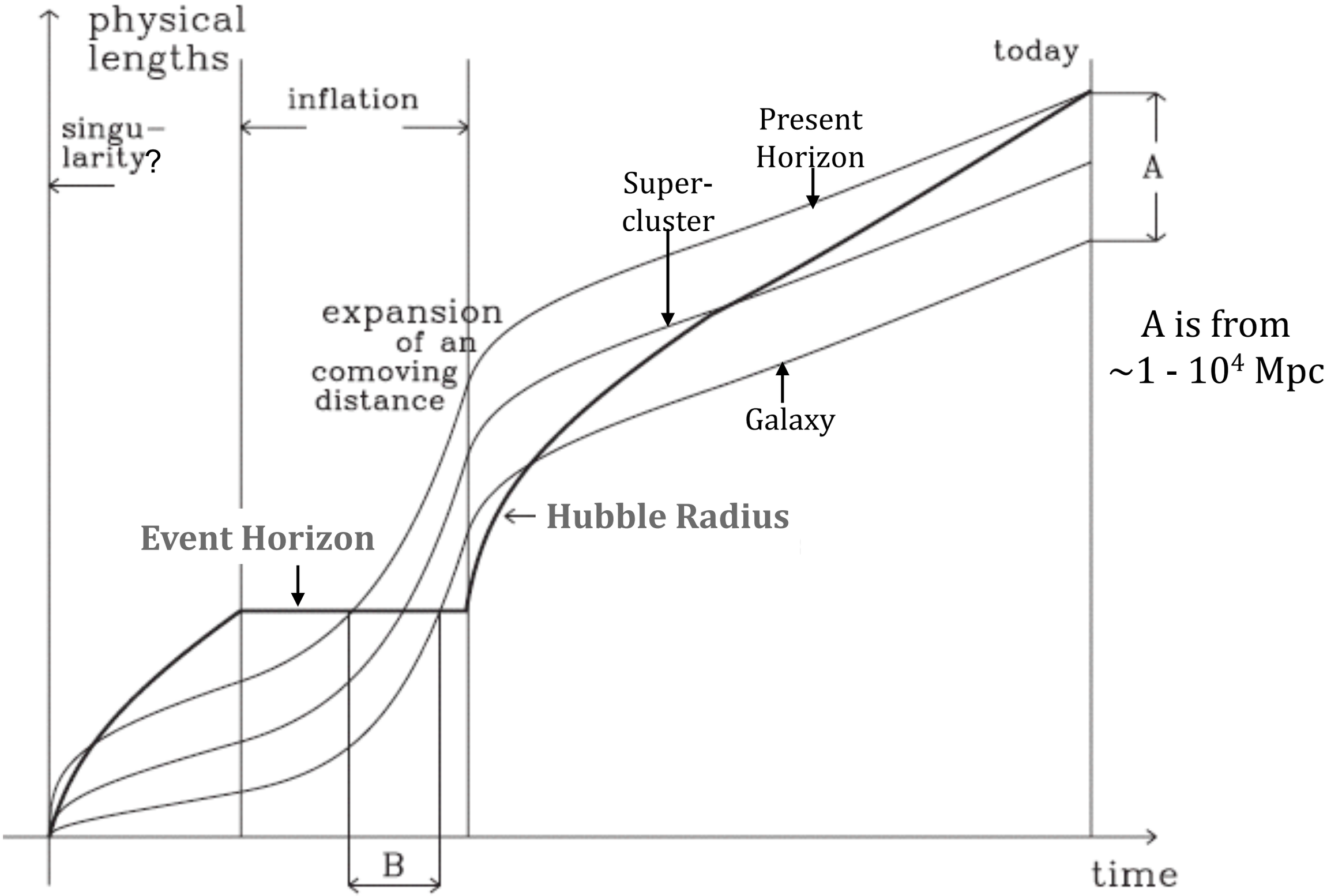
Caitlin Johnson – Observing Plan for ACTs to Constrain the Long-Wavelength EBL

Tanmayi Sai – Shapes of Simulated Galaxies

Vivian Tang – Studying Galaxy Evolution by Comparing Simulations and Observations

Please prepare a ~30 minute talk with slides.

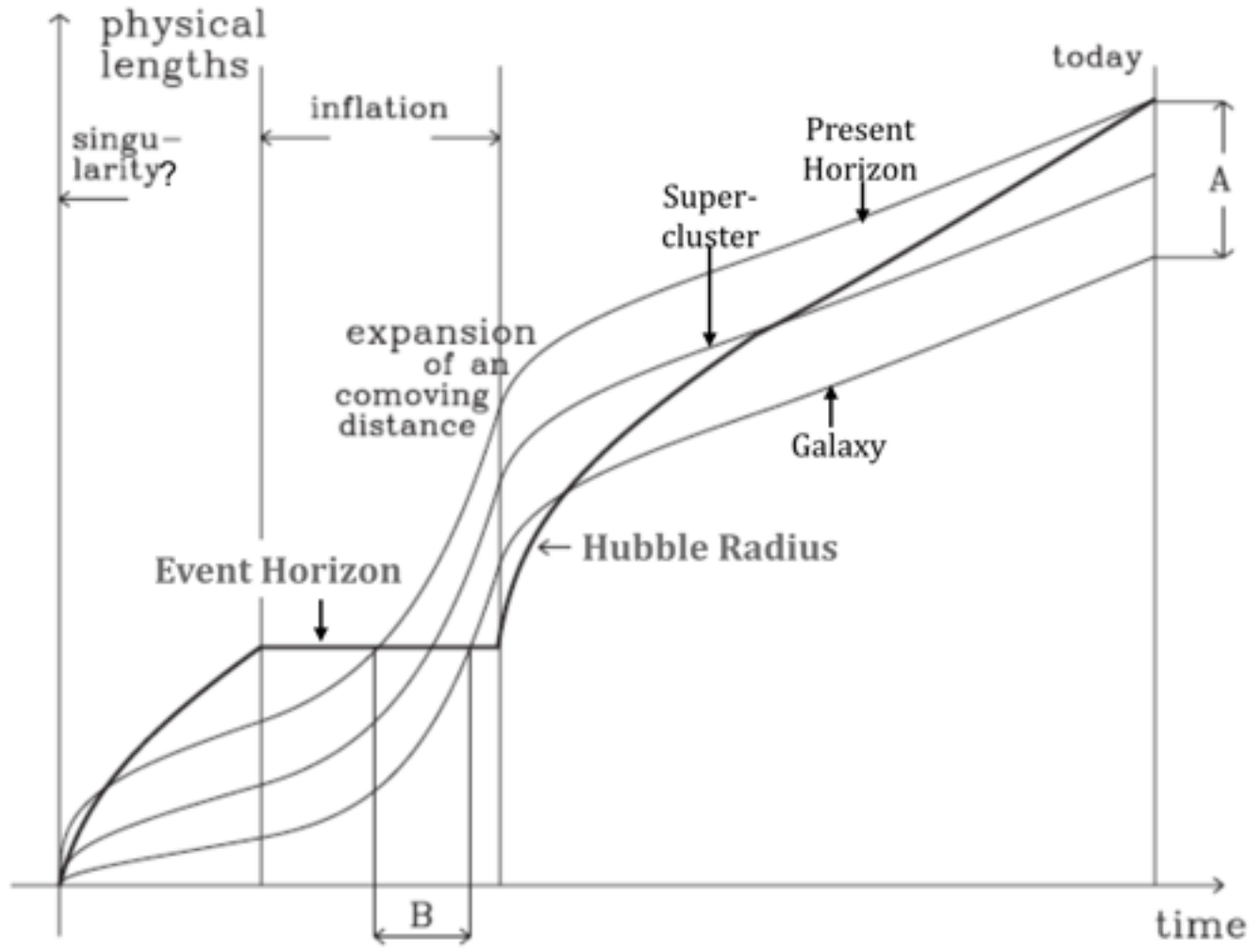
Dierck-Ekkehard Liebscher, *Cosmology* (Springer, 2005)



The evolution of the scales of perturbations. The larger scales overtake the Hubble radius at an early time and fall below it again later. They measure the inflation at an earlier time than do the smaller scales, which overtake the Hubble radius during inflation later and fall below it again earlier. The region *A* of scales that are accessible to evaluation today corresponds to a time span *B* of the inflation and related values of the inflaton field; for this time span, we can tell something – at least in principle – about the potential of the inflaton.

Fluctuations in Inflation

LOFI: The last scales to cross outside the horizon in Inflation are the first to cross inside in the subsequent Friedmann-Robertson-Walker phase. If the present cosmic horizon, which encompasses a mass of about $10^{22} M_\odot$, crossed outside the de Sitter event horizon at (say) 60 e-folds before the end of Inflation, galaxies (about $10^{12} M_\odot$ in mass) crossed about 52 e-folds before the end, and any horizon mass M_H crossed at $N = 60 + (1/3)\ln(M_H/10^{22} M_\odot)$.



There are quantum fluctuations in a de Sitter universe corresponding to the Hawking radiation temperature $T_H = H/2\pi$, so $< (\Delta\phi)^2 >= (H/2\pi)^2$. (Guth & Pi 1982 got the same answer using the de Sitter propagator.) This leads to density fluctuations $\delta\rho = V'\Delta\phi = -3H\dot{\phi}\Delta\phi$, where in the last step we used the slow-roll approximation $V' \equiv dV/d\phi = -3H\dot{\phi}$.

Thus in Inflation, $\rho = \dot{\phi}^2/2 + V(\phi) + \rho_{rad}$ and $p = \dot{\phi}^2/2 - V(\phi) + \rho_{rad}/3$, and we expect that ρ_{rad} will be negligible in Inflation. Note then that $\rho + p = \dot{\phi}^2$.

Bardeen's gauge invariant parameter $\zeta = \delta\rho/(\rho + p)$ is constant outside the horizon, so we equate the value of ζ as a given scale goes outside the horizon in Inflation with the value when it comes back inside:

$$\text{Inflation}: \quad \zeta = \frac{\delta\rho}{\rho+p} = \frac{\delta\rho}{\dot{\phi}^2} = \frac{V'\Delta\phi}{\dot{\phi}^2} = \frac{-3H\dot{\phi}\Delta\phi}{\dot{\phi}^2} = \frac{-3H^2}{2\pi\dot{\phi}}$$

$$\text{FRW}: \quad \zeta = \frac{\delta\rho}{\rho+p} = \frac{\delta\rho}{(4/3)\rho}$$

When the scale comes back inside the horizon in the Friedmann-Robertson-Walker phase, $\dot{\phi}^2$ and $V(\phi)$ are negligible and $\rho + p = (4/3)\rho$. Then

$$\frac{\delta\rho}{\rho} = \frac{4}{3}\zeta = \frac{-2H^2}{\pi\dot{\phi}} = \frac{6H^3}{\pi V'(\phi)} = \frac{6(8\pi V/3m_{\rm Pl}^2)^{3/2}}{\pi V'(\phi)},$$

where the next-to-last expression assumes that the slow-roll equation $V' \equiv dV/d\phi = -3H\dot{\phi}$ applies when the scale crossed outside the horizon in Inflation.

Eternal Inflation

Vilenkin (1983) and Linde (1986, 1990) pointed out that if one extrapolates inflation backward to try to imagine what might have preceeded it, in many versions of inflation the answer is "eternal inflation": in most of the volume of the universe inflation is still happening, and our part of the expanding universe (a region encompassing far more than our entire cosmic horizon) arose from a tiny part of such a region. To see how eternal inflation works, consider the simple chaotic model with $V(\phi) = (m^2/2)\phi^2$. During the de Sitter Hubble time H^{-1}, where as usual $H^2 = (8\pi G/3)V$, the slow rolling of ϕ down the potential will reduce it by

$$\Delta\phi = \dot{\phi}\Delta t = -\frac{V'}{3H}\Delta t = \frac{m_{pl}^2}{4\pi\phi}. \tag{1.7}$$

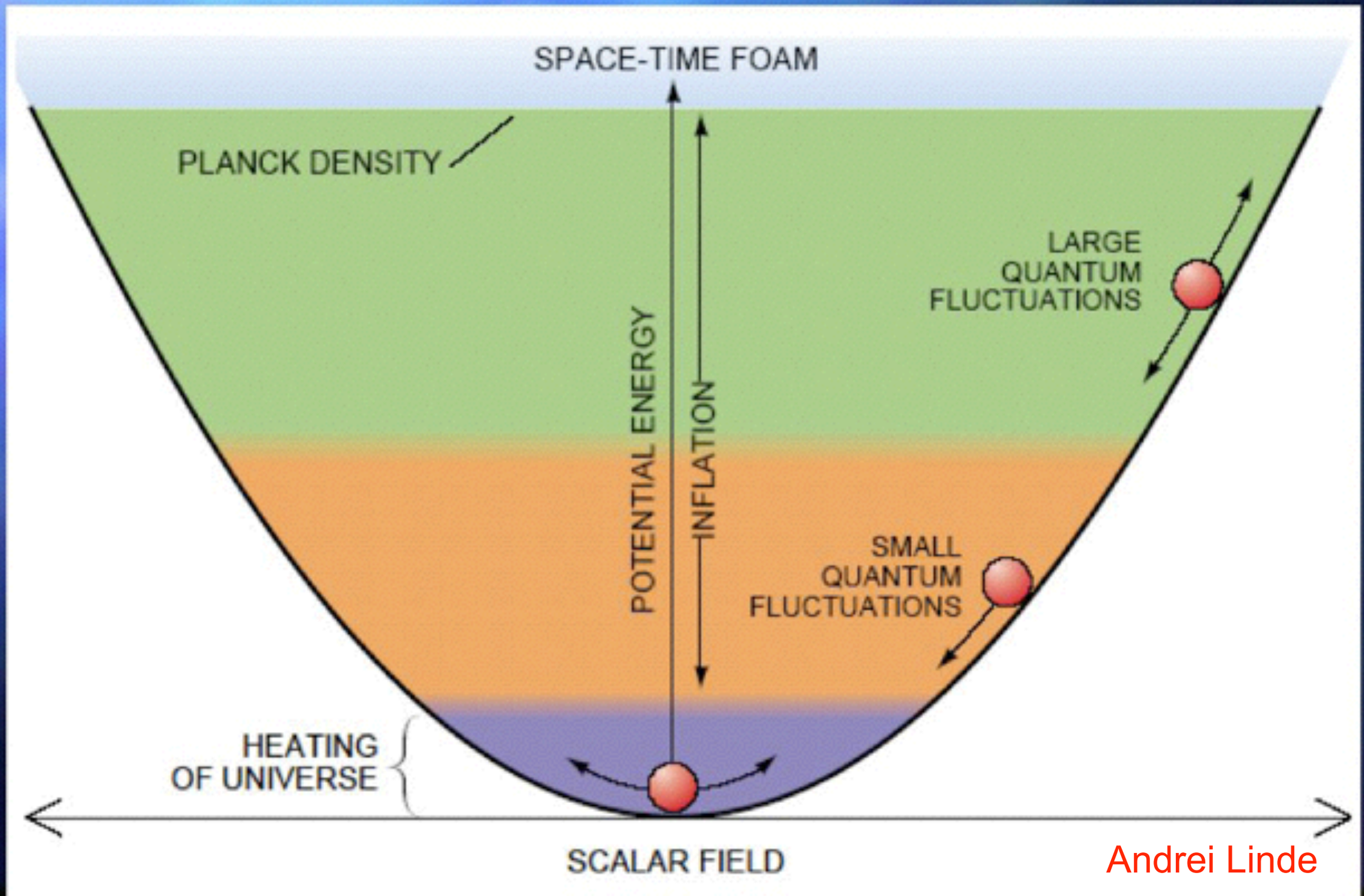
Here m_{Pl} is the Planck mass ($m_{Planck} = 1/G^{1/2}$). But there will also be quantum fluctuations that will change ϕ up or down by

$$\delta\phi = \frac{H}{2\pi} = \frac{m\phi}{\sqrt{3\pi}m_{Pl}} \tag{1.8}$$

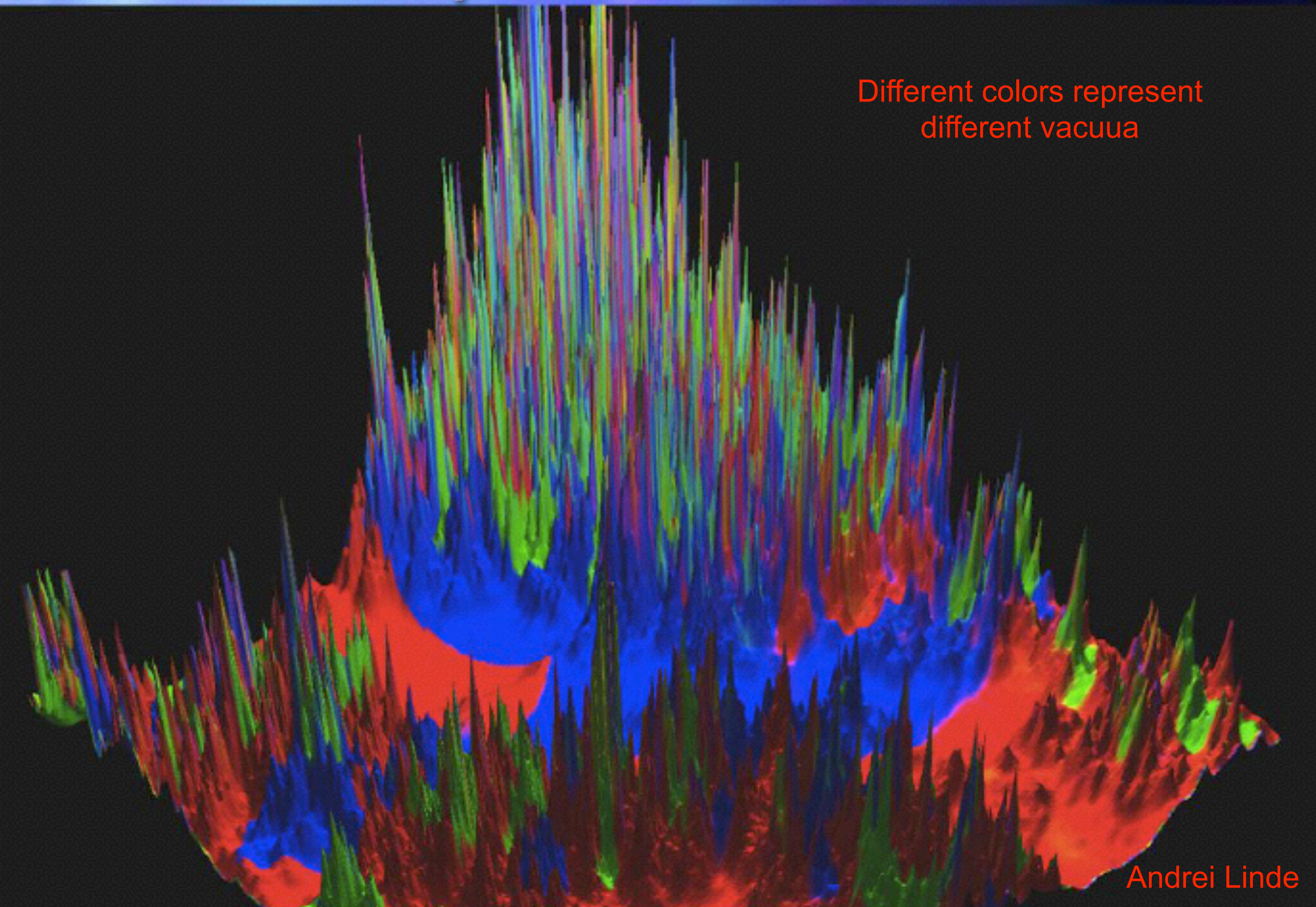
These will be equal for $\phi_* = m_{pl}^{3/2}/2m^{1/2}$, $V(\phi_*) = (m/8m_{Pl})m_{Pl}^4$. If $\phi \gtrsim \phi_*$, *positive quantum fluctuations dominate* the evolution: after $\Delta t \sim H^{-1}$, an initial region becomes $\sim e^3$ regions of size $\sim H^{-1}$, in half of which ϕ increases to $\phi + \delta\phi$. Since $H \propto \phi$, this drives inflation faster in these regions.

Inflation as a theory of a harmonic oscillator

$$V(\phi) = \frac{m^2}{2}\phi^2$$



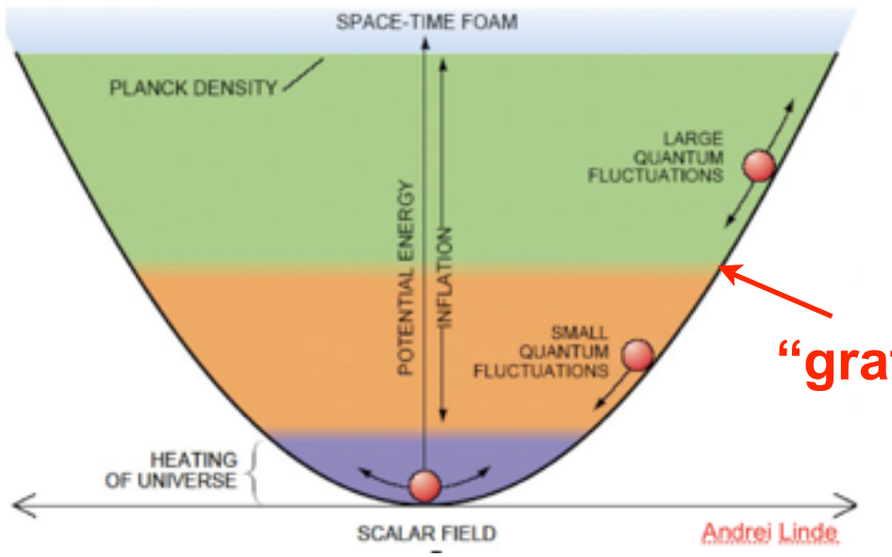
Landscape of eternal inflation



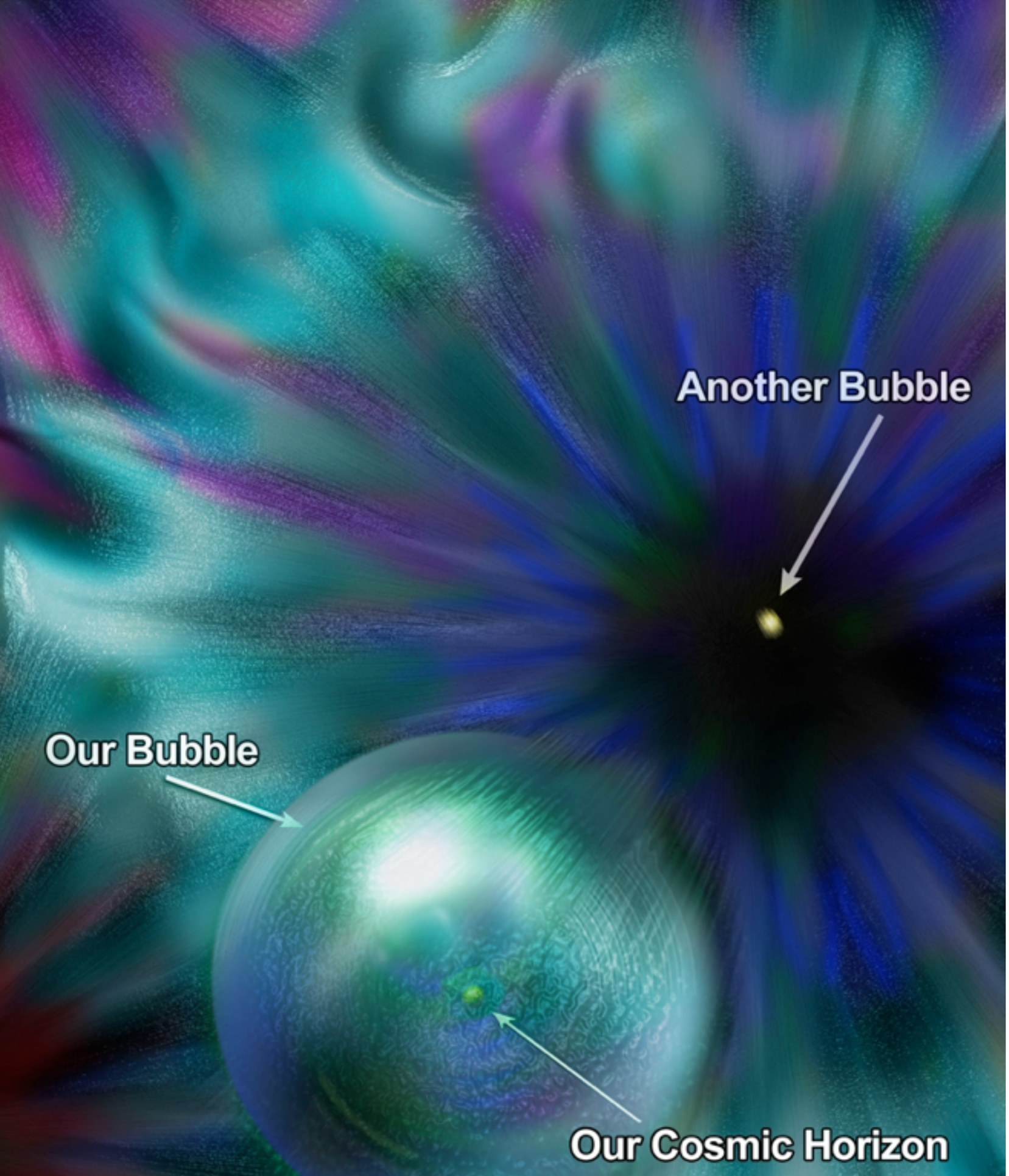
THE COSMIC LAS VEGAS

Coins constantly flip. Heads, and the coin is twice the size and there are two of them. Tails, and a coin is half the size.

Consider a coin that has a run of tails. It becomes so small it can pass through the grating on the floor.

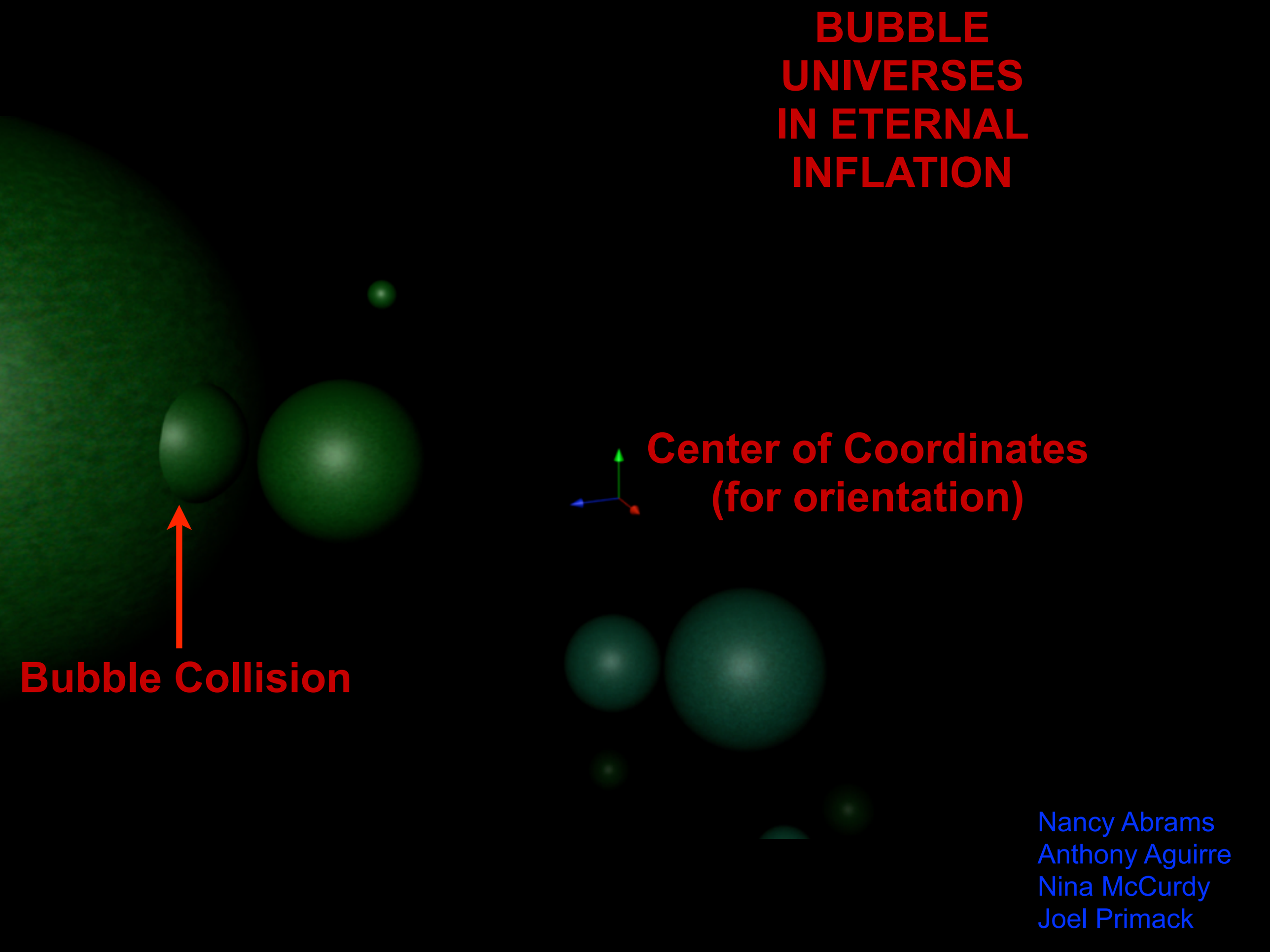


“grating”



OUR COSMIC BUBBLE IN ETERNAL INFLATION

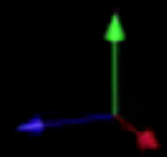
BUBBLE UNIVERSES IN ETERNAL INFLATION



Nancy Abrams
Anthony Aguirre
Nina McCurdy
Joel Primack

Expanding Bubbles
Getting Dimmer
Are Receding

BUBBLE
UNIVERSES
IN ETERNAL
INFLATION



Supersymmetric Inflation

When Pagels and I (1982) first suggested that the lightest supersymmetric partner particle (LSP), stable because of R-parity, might be the dark matter particle, that particle was the gravitino in the early version of supersymmetry then in fashion. Weinberg (1982) immediately pointed out that if the gravitino were not the LSP, it could be a source of real trouble because of its long lifetime $\sim M_{\rm Pl}^2/m_{3/2}^3 \sim (m_{3/2}/{\rm TeV})^{-3}10^3$ s, a consequence of its gravitational-strength coupling to other fields. Subsequently, it was realized that supersymmetric theories can naturally solve the gauge hierarchy problem, explaining why the electroweak scale $M_{\rm EW} \sim 10^2$ GeV is so much smaller than the GUT or Planck scales. In this version of supersymmetry, which has now become the standard one, the gravitino mass will typically be $m_{3/2} \sim$ TeV; and the late decay of even a relatively small number of such massive particles can wreck BBN and/or the thermal spectrum of the CBR. The only way to prevent this is to make sure that the reheating temperature after inflation is sufficiently low: $T_{\rm RH} \lesssim 2 \times 10^9$ GeV (for $m_{3/2} =$ TeV) (Ellis, Kim, & Nanopoulos 1984, Ellis et al. 1992).

Inflation Summary

The key features of all inflation scenarios are a period of superluminal expansion, followed by ("re-")heating which converts the energy stored in the inflaton field (for example) into the thermal energy of the hot big bang.

Inflation is *generic*: it fits into many versions of particle physics, and it can even be made rather natural in modern supersymmetric theories as we have seen. The simplest models have inflated away all relics of any pre-inflationary era and result in a flat universe after inflation, i.e., $\Omega = 1$ (or more generally $\Omega_0 + \Omega_\Lambda = 1$). Inflation also produces scalar (density) fluctuations that have a primordial spectrum

$$\left(\frac{\delta\rho}{\rho}\right)^2 \sim \left(\frac{V^{3/2}}{m_{Pl}^3 V'}\right)^2 \propto k^{n_p}, \tag{1.12}$$

where V is the inflaton potential and n_p is the primordial spectral index, which is expected to be near unity (near-Zel'dovich spectrum). Inflation also produces tensor (gravity wave) fluctuations, with spectrum

$$P_t(k) \sim \left(\frac{V}{m_{Pl}}\right)^2 \propto k^{n_t}, \tag{1.13}$$

where the tensor spectral index $n_t \approx (1 - n_p)$ in many models.

The quantity $(1-n_p)$ is often called the "tilt" of the spectrum; the larger the tilt, the more fluctuations on small spatial scales (corresponding to large k) are suppressed compared to those on larger scales. The scalar and tensor waves are generated by independent quantum fluctuations during inflation, and so their contributions to the CMB temperature fluctuations add in quadrature. The ratio of these contributions to the quadrupole anisotropy amplitude Q is often called $T/S \equiv Q_t^2/Q_s^2$; thus the primordial scalar fluctuation power is decreased by the ratio $1/(1+T/S)$ for the same COBE normalization, compared to the situation with no gravity waves ($T=0$). In power-law inflation, $T/S = 7(1-n_p)$. This is an approximate equality in other popular inflation models such as chaotic inflation with $V(\phi) = m^2\phi^2$ or $\lambda\phi^4$. But note that the tensor wave amplitude is just the inflaton potential during inflation divided by the Planck mass, so the gravity wave contribution is negligible in theories like the supersymmetric model discussed above in which inflation occurs at an energy scale far below m_{Pl}. Because gravity waves just redshift after they come inside the horizon, the tensor contributions to CMB anisotropies corresponding to angular wavenumbers $\ell \gg 20$, which came inside the horizon long ago, are strongly suppressed compared to those of scalar fluctuations.

Joel Primack, in *Formation of Structure in the Universe*, ed. Dekel & Ostriker (Cambridge Univ Press, 1999)

Basic Predictions of Inflation

1. **Flat universe.** This is perhaps the most fundamental prediction of inflation. Through the Friedmann equation it implies that the total energy density is always equal to the critical energy density; it does not however predict the form (or forms) that the critical density takes on today or at any earlier or later epoch.

2. **Nearly scale-invariant spectrum of Gaussian density perturbations.** These density perturbations (scalar metric perturbations) arise from quantum-mechanical fuctuations in the field that drives inflation; they begin on very tiny scales (of the order of 10^{-23} cm, and are stretched to astrophysical size by the tremendous growth of the scale factor during inflation (factor of e^{60} or greater). Scale invariant refers to the fact that the fuctuations in the gravitational potential are independent of length scale; or equivalently that the horizon-crossing amplitudes of the density perturbations are independent of length scale. While the shape of the spectrum of density perturbations is common to all models, the overall amplitude is model dependent. Achieving density perturbations that are consistent with the observed anisotropy of the CBR and large enough to produce the structure seen in the Universe today requires a horizon crossing amplitude of around 2×10^{-5}.

3. **Nearly scale-invariant spectrum of gravitational waves**, from quantum-mechanical fluctuations in the metric itself . These can be detected as CMB “B-mode” polarization, or using special gravity wave detectors such as LIGO and LISA.

Useful Formulas

Density Fluctuations from Inflation

The relationship between the inflationary potential and the power spectrum of density perturbations today ($P(k) \equiv \langle |\delta_k|^2 \rangle$) is given by

Transfer function

Power Spectrum
$$P(k) = \frac{1024\pi^3}{75} \frac{k}{H_0^4} \frac{V_*^3}{{m_{\rm Pl}}^6 V_*'^2} \left(\frac{k}{k_*}\right)^{n-1} T^2(k)$$

Tilt
$$n_{\rm S} - 1 = -\frac{1}{8\pi}\left(\frac{m_{\rm Pl} V_*'}{V_*}\right)^2 + \frac{m_{\rm Pl}}{4\pi}\left(\frac{m_{\rm Pl} V_*'}{V_*}\right)'$$

generally nonzero, ≈ 0.04 according to WMAP & Planck

Running Tilt
$$\frac{dn}{d\ln k} = -\frac{1}{32\pi^2}\left(\frac{{m_{\rm Pl}}^3 V_*'''}{V_*}\right)\left(\frac{m_{\rm Pl} V_*'}{V_*}\right) + \frac{1}{8\pi^2}\left(\frac{{m_{\rm Pl}}^2 V_*''}{V_*}\right)\left(\frac{m_{\rm Pl} V_*'}{V_*}\right)^2 - \frac{3}{32\pi^2}\left(m_{\rm Pl}\frac{V_*'}{V_*}\right)^4$$

$$T(q) = \frac{\ln(1+2.34q)/2.34q}{\left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{1/4}}, \qquad (4)$$

where $V(\phi)$ is the inflationary potential, prime denotes $d/d\phi$, V_* is the value of the scalar potential when the scale k_* crossed outside the horizon during inflation, $T(k)$ is the transfer function which accounts for the evolution of the mode k from horizon crossing until the present, $q = k/h\Gamma$, and $\Gamma \simeq \Omega_M h$ is the "shape" parameter . The fitting formula (4) isn't accurate enough for precision work; instead, use the website http://camb.info/ .

Useful Formulas

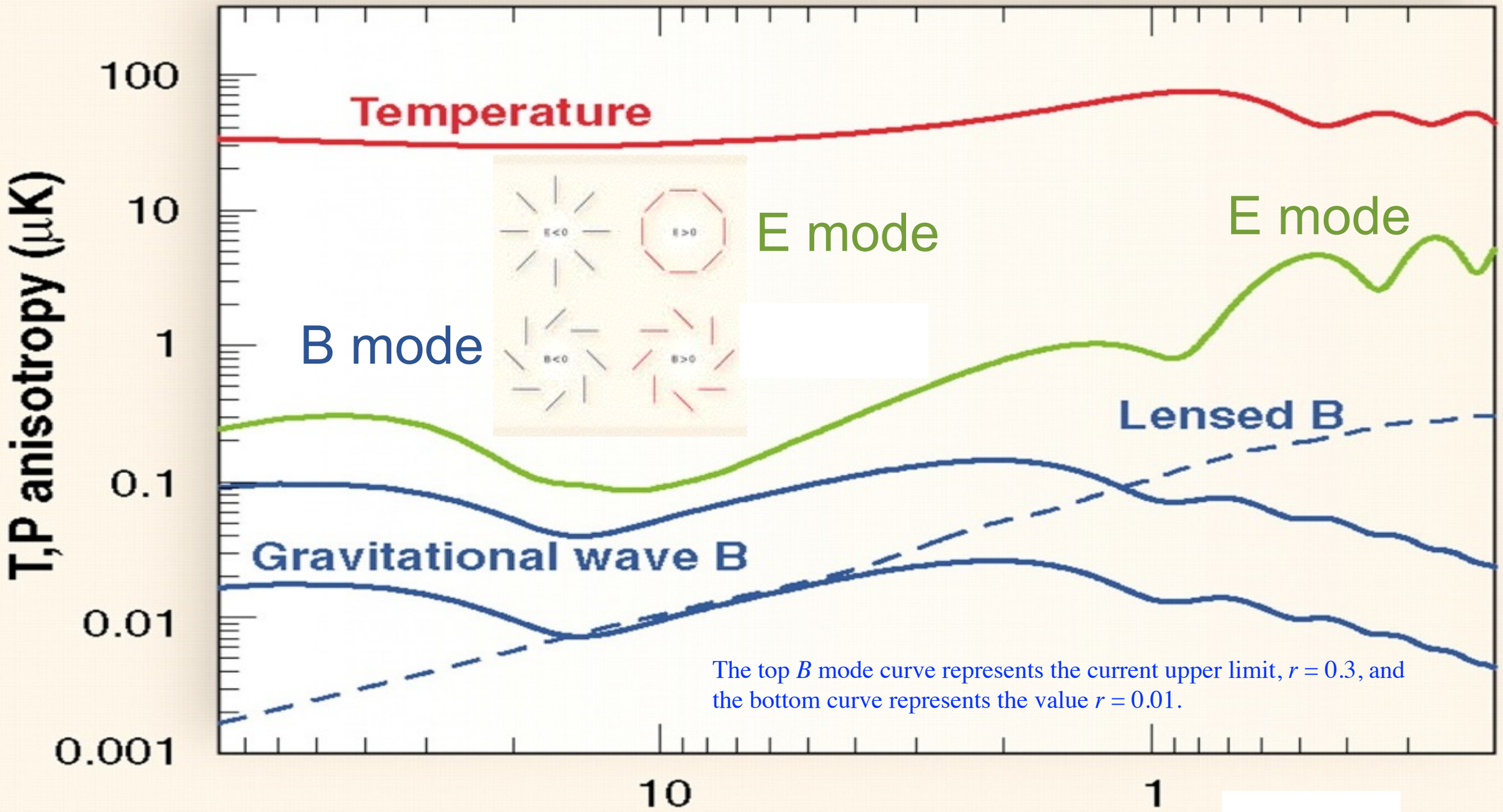
Gravity Waves from Inflation

Unlike the scalar perturbations, which must have an amplitude of around 10^{-5} to seed structure formation, there is an upper but no lower limit on the amplitude of the tensor perturbations. They can be characterized by their power spectrum today

$$
\begin{aligned}
P_T(k) &\equiv \langle |h_k|^2 \rangle = \frac{8}{3\pi} \frac{V_*}{{m_{\rm Pl}}^4} \left(\frac{k}{k_*}\right)^{n_T - 3} T_T^2(k) \\
n_T &= -\frac{1}{8\pi} \left(\frac{m_{\rm Pl} V_*'}{V_*}\right)^2 \\
\frac{dn_T}{d\ln k} &= \frac{1}{32\pi^2} \left(\frac{{m_{\rm Pl}}^2 V''}{V}\right) \left(\frac{m_{\rm Pl} V'}{V}\right)^2 - \frac{1}{32\pi^2} \left(\frac{m_{\rm Pl} V'}{V}\right)^4 = -n_T[(n-1) - n_T] \\
T_T(k) &\simeq \left[1 + \frac{4}{3}\frac{k}{k_{\rm EQ}} + \frac{5}{2}\left(\frac{k}{k_{\rm EQ}}\right)^2\right]^{1/2}, \qquad (11)
\end{aligned}
$$

where $T_T(k)$ is the transfer function for gravity waves and describes the evolution of mode k from horizon crossing until the present, $k_{\rm EQ} = 6.22 \times 10^{-2}\,{\rm Mpc}^{-1}\,(\Omega_M h^2/\sqrt{g_*/3.36})$ is the scale that crossed the horizon at matter-radiation equality, Ω_M is the fraction of critical density in matter, and g_* counts the effective number of relativistic degrees of freedom (3.36 for photons and three light neutrino species). The quantity $k^{3/2}|h_k|/\sqrt{2\pi^2}$ corresponds to the dimensionless strain (metric perturbation) on length scale $\lambda = 2\pi/k$.

Root mean square fluctuations in temperature (T) and polarization (E and B modes) of the CMB predicted by inflation.



$$V^{1/4} = 1.06 \times 10^{16}\,\mathrm{GeV}\left(\frac{r}{0.01}\right)$$

$$r \equiv \frac{P_t}{P_s}$$

L M Krauss, S. Dodelson, S. Meyer Science 2010;328:989-992

Observational Status of Inflation

I. The predictions of inflation are right:

(i) the universe is flat with a critical density
(ii) superhorizon fluctuations
(iii) density perturbation spectrum nearly scale invariant: $P(k) = Ak^n$, $n \approx 1$
(iv) Single slow-roll field models vindicated: Gaussian perturbations, not much running of spectral index

- If primordial fluctuations are Gaussian distributed, then they are completely characterized by their two-point function $\xi(r)$, or equivalently by the power spectrum. All odd-point functions are zero.
- If nonGaussian, there is additional info in the higher order correlation functions
- The lowest order statistic that can differentiate is the 3-point function, or bispectrum in Fourier space: $\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_\Phi(k_1,k_2,k_3)$. Here $B_\Phi(k_1,k_2,k_3) = f_{\rm NL}F(k_1,k_2,k_3)$. The quantity $f_{\rm NL}$ is known as the nonlinearity parameter. Planck data: $f_{\rm NL} = 2.7\pm5.8$ -- small!

II. Data differentiate between models

-- each model makes specific predictions for density perturbations and gravity modes
-- WMAP and Planck rule out many models (see graph on next page, from Planck 2013 Results XXII, http://arxiv.org/abs/1303.5082).

Observational Status of Inflation

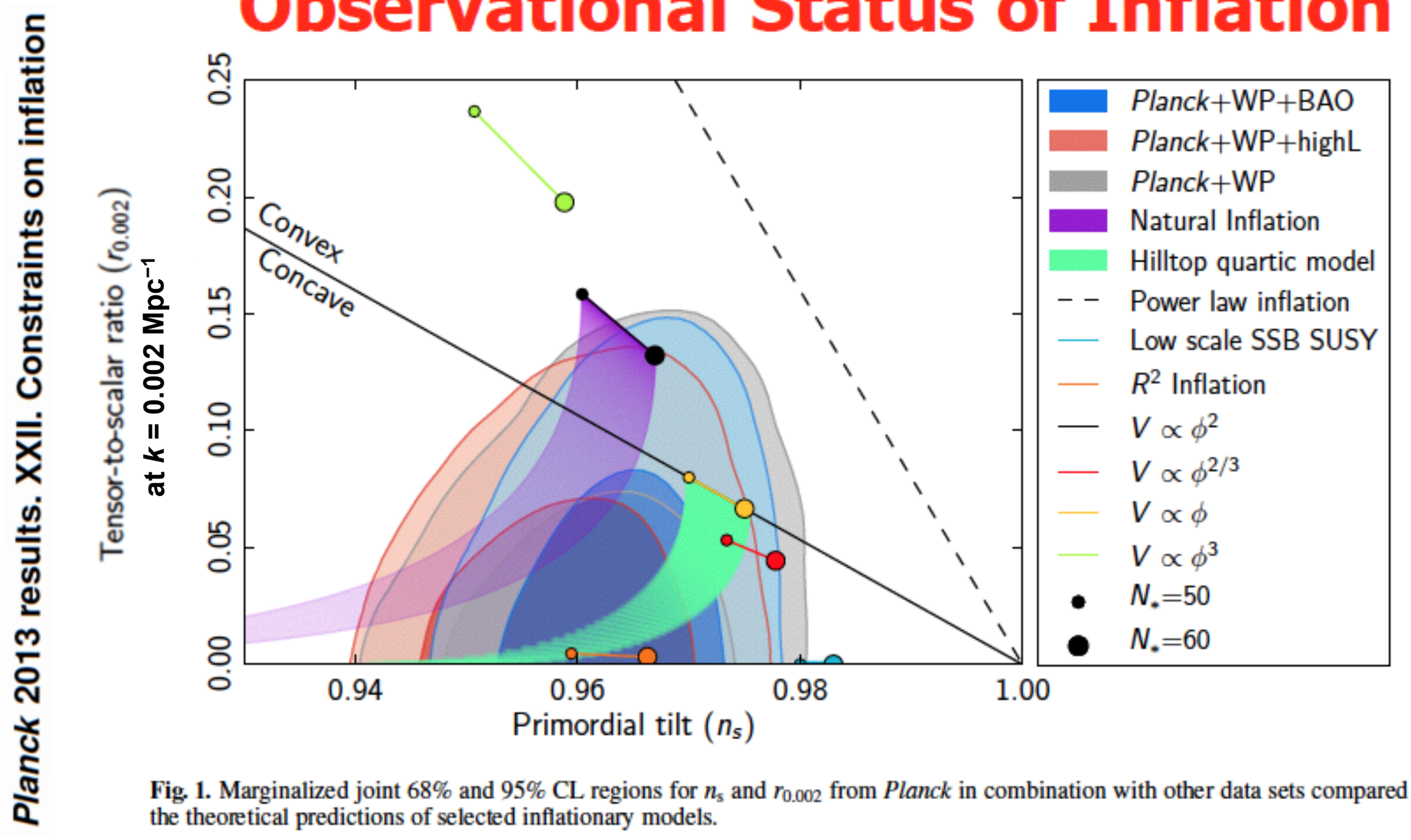
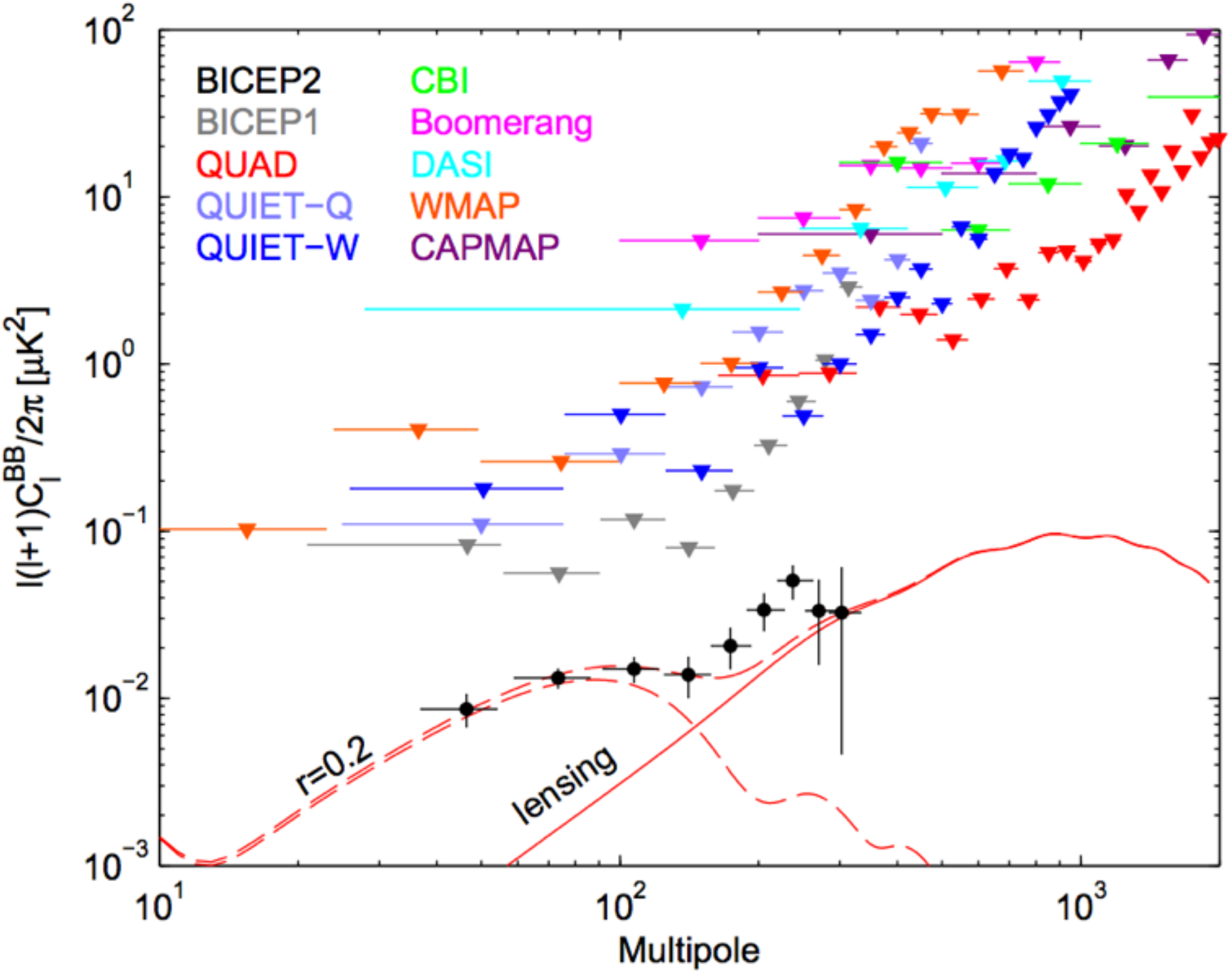


Fig. 1. Marginalized joint 68% and 95% CL regions for n_s and $r_{0.002}$ from *Planck* in combination with other data sets compared to the theoretical predictions of selected inflationary models.

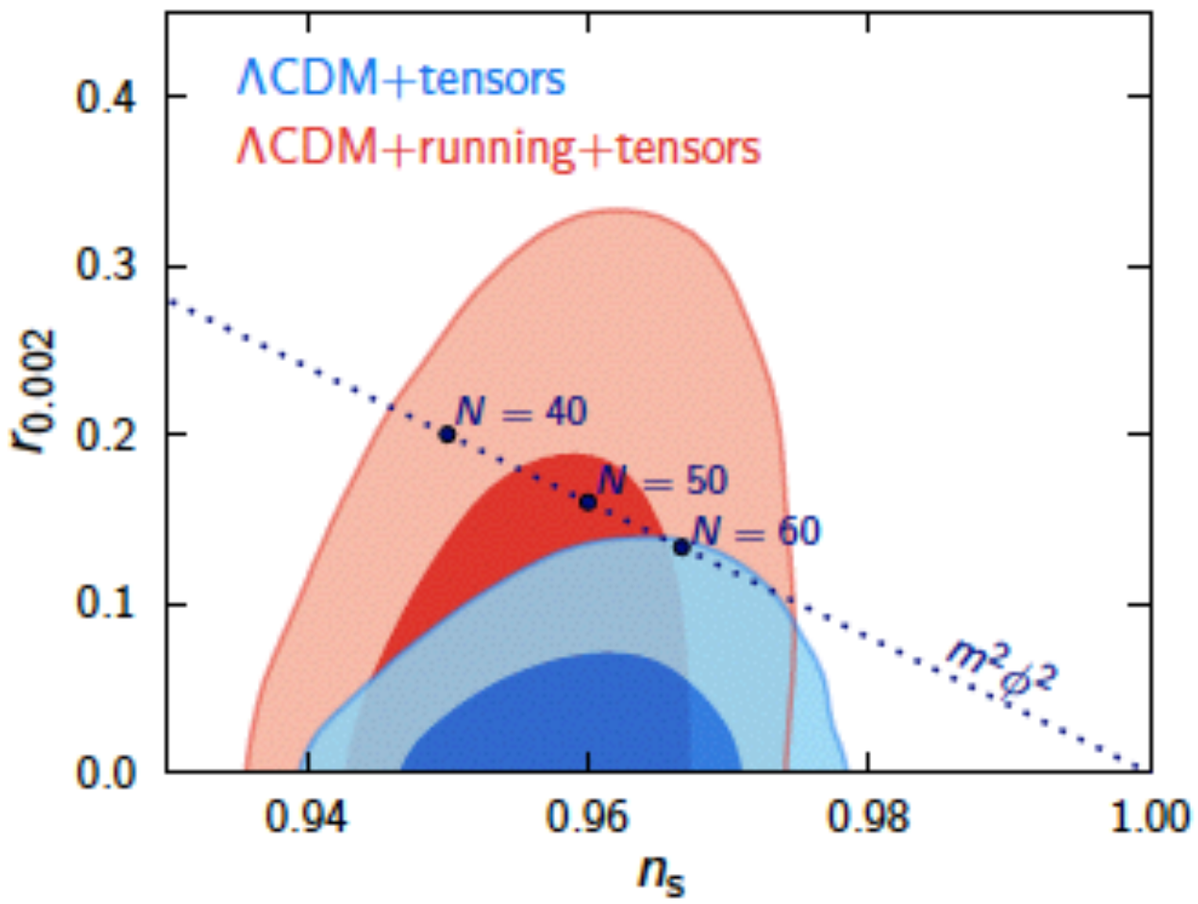
CONCLUSIONS: We find that standard slow-roll single field inflation is compatible with the Planck data. Planck in combination with WMAP 9-year large angular scale polarization (WP) yields Ω_K −0.006 ±0.018 at 95%CL by combining temperature and lensing information (Planck Collaboration XVI, 2013; Planck Collaboration XVII, 2013). The bispectral non-Gaussianity parameter f_{NL} measured by Planck is consistent with zero (Planck Collaboration XXIV, 2013). These results are compatible with zero spatial curvature and a small value of f_{NL}, as predicted in the simplest slow-roll inflationary models. Planck+WP data give n_s =0.9603±0.0073 (and n_s = 0.9629±0.0057 when combined with BAO). The 95% CL bound on the tensor-to-scalar ratio is r 0.002 < 0.12; this implies an upper limit for the inflation energy scale of 1.9×10^{16} GeV.

BICEP2 I: DETECTION OF B-mode POLARIZATION AT DEGREE ANGULAR SCALES

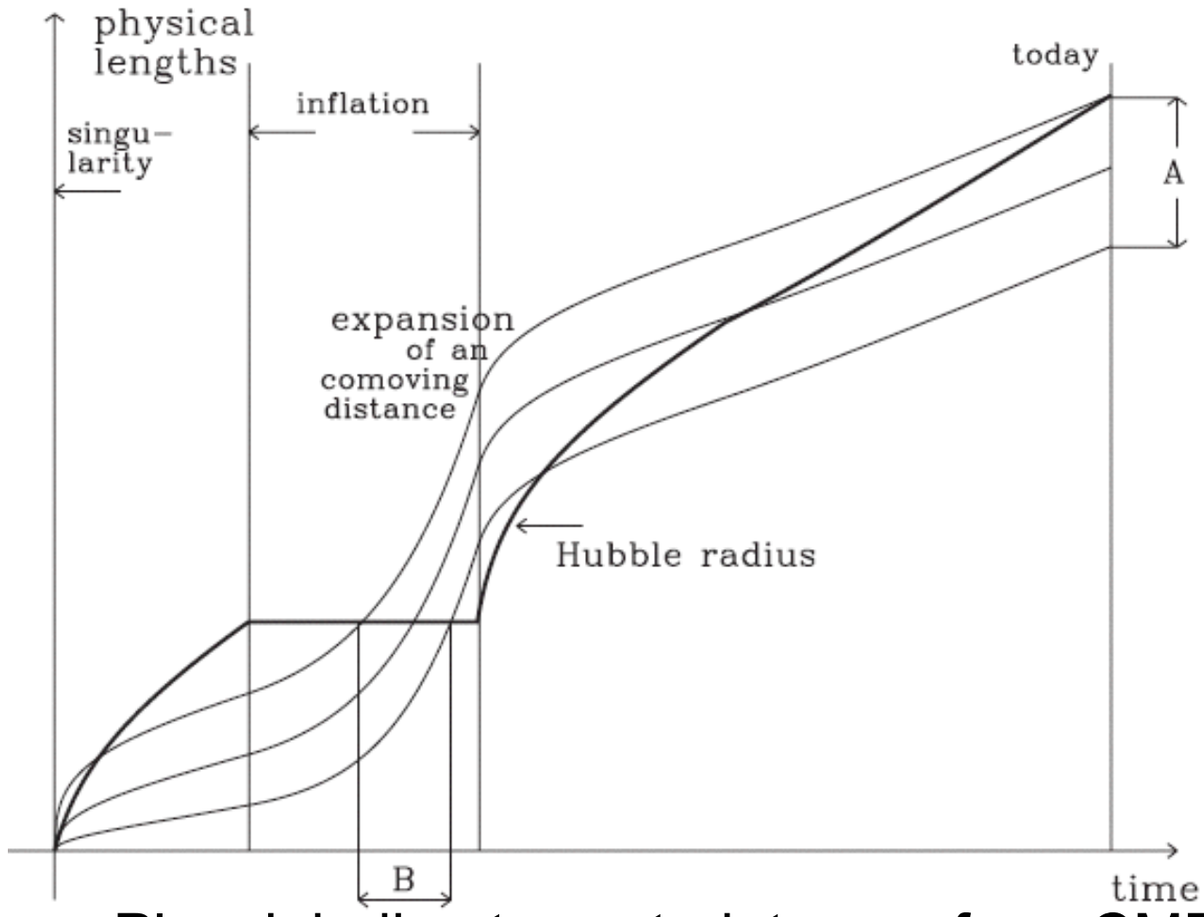


BICEP2 BB auto spectra and 95% upper limits from several previous experiments. The curves show the theory expectations for Tensor/Scalar = r = 0.2 and lensed-ΛCDM.

BICEP2 I: DETECTION OF B-mode POLARIZATION AT DEGREE ANGULAR SCALES



Planck Collaboration XVI (2013) Figure 23: Posterior distribution for n_s for the ΛCDM model with tensors (blue) compared to the posterior when a tensor component and running scalar spectral index are added to the model (red). The dotted line shows the relation between *r* and n_s for a *V*(φ) inflaton potential where *N* is the number of inflationary *e*-foldings.



Planck indirect constraints on r from CMB temperature spectrum measurements relax in the context of various model extensions. Shown here, following Planck Collaboration XVI (2013) Figure 23, with tensors and running of the scalar spectral index added to the base ΛCDM model. The contours show the resulting 68% and 95% confidence regions for r and the scalar spectral index n_s when also allowing running.

Post-Inflation

Baryogenesis: generation of excess of baryon (and lepton) number compared to anti-baryon (and anti-lepton) number. In order to create the observed baryon number today

$$\frac{n_B}{n_\gamma} = \left(6.1\,^{+0.3}_{-0.2}\right) \times 10^{-10}$$

it is only necessary to create an excess of about 1 quark and lepton for every ~10^9 quarks+antiquarks and leptons +antileptons.

Other things that might happen Post-Inflation:

Breaking of Pecci-Quinn symmetry so that the observable universe is composed of many PQ domains.

Formation of cosmic topological defects if their amplitude is small enough not to violate cosmological bounds.

Post-Inflation

Baryogenesis: generation of excess of baryon (and lepton) number compared to anti-baryon (and anti-lepton) number. In order to create the observed baryon number today

$$\frac{n_B}{n_\gamma} = \left(6.1\,^{+0.3}_{-0.2}\right) \times 10^{-10}$$

it is only necessary to create an excess of about 1 quark and lepton for every ~10^9 quarks+antiquarks and leptons +antileptons.

Other things that might happen Post-Inflation:

Breaking of Pecci-Quinn symmetry so that the observable universe is composed of many PQ domains.

Formation of cosmic topological defects if their amplitude is small enough not to violate cosmological bounds.

Baryogenesis

There is good evidence that there are **no large regions of antimatter** (Cohen, De Rujula, and Glashow, 1998). It was **Andrei Sakharov** (1967) who first suggested that the baryon density might not represent some sort of initial condition, but might be understandable in terms of microphysical laws. He listed three ingredients to such an understanding:

1. **Baryon number violation** must occur in the fundamental laws. At very early times, if baryon number violating interactions were in equilibrium, then the universe can be said to have "started" with zero baryon number. Starting with zero baryon number, baryon number violating interactions are obviously necessary if the universe is to end up with a non-zero asymmetry. As we will see, apart from the philosophical appeal of these ideas, the success of inflationary theory suggests that, shortly after the big bang, the baryon number was essentially zero.

2. **CP-violation**: If CP (the product of charge conjugation and parity) is conserved, every reaction which produces a particle will be accompanied by a reaction which produces its antiparticle at precisely the same rate, so no baryon number can be generated.

3. **Departure from Thermal Equilibrium** (An Arrow of Time): The universe, for much of its history, was very nearly in thermal equilibrium. The spectrum of the CMBR is the most perfect blackbody spectrum measured in nature. So the universe was certainly in thermal equilibrium 10^5 years after the big bang. The success of the theory of big bang nucleosynthesis (BBN) provides strong evidence that the universe was in equilibrium two-three minutes after the big bang. But if, through its early history, the universe was in thermal equilibrium, then even B and CP violating interactions could not produce a net asymmetry. One way to understand this is to recall that the CPT theorem assures strict equality of particle and antiparticle masses, so at thermal equilibrium, the densities of particles and antiparticles are equal. More precisely, since B is odd under CPT, its thermal average vanishes in an equilibrium situation. This can be generalized by saying that the universe must have an arrow of time.

Several mechanisms have been proposed to understand the baryon asymmetry:

1. **GUT Baryogenesis**. Grand Unified Theories unify the gauge interactions of the strong, weak and electromagnetic interactions in a single gauge group. They inevitably violate baryon number, and they have heavy particles, with mass of order $M_{GUT} \approx 10^{16}$ GeV, whose decays can provide a departure from equilibrium. The main objections to this possibility come from issues associated with inflation. While there does not exist a compelling microphysical model for inflation, in most models, the temperature of the universe after reheating is well below M_{GUT}. But even if it were very large, there would be another problem. Successful unification requires supersymmetry, which implies that the graviton has a spin-3/2 partner, called the gravitino. In most models for supersymmetry breaking, these particles have masses of order TeV, and are very long lived. Even though these particles are weakly interacting, too many gravitinos are produced unless the reheating temperature is well below the unification scale -- too low for GUT baryogenesis to occur.

2. **Electroweak baryogenesis**. The Standard Model satisfies all of the conditions for baryogenesis, but any baryon asymmetry produced is far too small to account for observations. In certain extensions of the Standard Model, it is possible to obtain an adequate asymmetry, but in most cases the allowed region of parameter space is very small.

3. **Leptogenesis**. The possibility that the weak interactions will convert some lepton number to baryon number means that if one produces a large lepton number at some stage, this will be processed into a net baryon and lepton number at the electroweak phase transition. The observation of neutrino masses makes this idea highly plausible. Many but not all of the relevant parameters can be directly measured.

4. **Production by coherent motion of scalar fields (the Affleck-Dine mechanism)**, which can be highly efficient, might well be operative if nature is supersymmetric.

Following Dine & Kusenko, RMP 2004.

1. GUT Baryogenesis. GUTs satisfy all three of Sakharov's conditions.

Baryon number (B) violation is a hallmark of these theories: they typically contain gauge bosons and other fields which mediate B violating interactions such as proton decay.

CP violation is inevitable; necessarily, any model contains at least the Kobayashi-Maskawa (KM) mechanism for violating CP, and typically there are many new couplings which can violate CP.

Departure from equilibrium is associated with the dynamics of the massive, B violating fields. Typically one assumes that these fields are in equilibrium at temperatures well above the grand unification scale. As the temperature becomes comparable to their mass, the production rates of these particles fall below their rates of decay. Careful calculations in these models often lead to baryon densities compatible with what we observe.

Example: SU(5) GUT. Treat all quarks and leptons as left-handed fields. In a single generation of quarks and leptons one has the quark doublet Q, the singlet u-bar and d-bar antiquarks (their antiparticles are the right-handed quarks), and the lepton doublet, *L*.

Then it is natural to identify th
fields in the 5-bar as follows:

$$\bar{5}_i = \begin{pmatrix} \bar{d} \\ \bar{d} \\ \bar{d} \\ e \\ \nu \end{pmatrix}$$

The remaining quarks and leptons (e- and e+) are in a 10 of SU(5).

The gauge fields are in the 24 (adjoint) representation:

Color SU(3) $T = \begin{pmatrix} \frac{\lambda^a}{2} & 0 \\ 0 & 0 \end{pmatrix}$ Weak SU(2) $T = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\sigma^i}{2} \end{pmatrix}$

The U(1) generator is

$$Y' = \frac{1}{\sqrt{60}} \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & -3 & \\ & & & & -3 \end{pmatrix}$$

SU(5) is a broken symmetry, and it can be broken by a scalar Higgs field proportional to Y'. The unbroken symmetries are generated by the operators that commute with Y', namely SU(3)xSU(2)xU(1). The vector bosons *X* that correspond to broken generators, for example

$$\left(\begin{array}{ccc:cc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hdashline 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

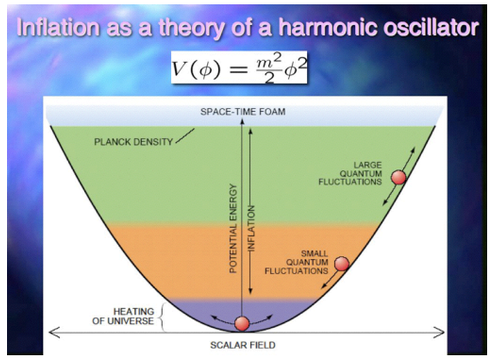
gain mass ~10^{16} GeV by this GUT Higgs mechanism.

The *X* bosons carry color and electroweak quantum numbers and mediate processes which violate baryon number. For example, there is a coupling of the X bosons to a d-bar quark and an electron.

In the GUT picture of baryogenesis, it is usually assumed that at temperatures well above the GUT scale, the universe was in thermal equilibrium. As the temperature drops below the mass of the X bosons, the reactions which produce the X bosons are not sufficiently rapid to maintain equilibrium. The decays of the X bosons violate baryon number; they also violate CP. So all three conditions are readily met: **B violation**, **CP violation**, and **departures from equilibrium**.

CPT requires that the total decay rate of X is the same as that of its antiparticle X-bar. But it does not require equality of the decays to particular final states (partial widths). So starting with equal numbers of X and X-bar particles, there can be a slight asymmetry between the processes

and



This can result in a slight excess of matter over anti-matter. But reheating to T >10^{16} GeV after inflation will overproduce gravitinos -- so GUT baryogenesis is now disfavored.

When Pagels and I (1982) first suggested that the lightest supersymmetric partner particle (LSP), stable because of R-parity, might be the dark matter particle, that particle was the gravitino in the early version of supersymmetry then in fashion. Weinberg (1982) immediately pointed out that if the gravitino were not the LSP, it could be a source or real trouble because of its long lifetime $\sim M_{\rm Pl}^2/m_{3/2}^3 \sim (m_{3/2}/{\rm TeV})^{-3}10^3$ s, a consequence of its gravitational-strength coupling to other fields. Subsequently, it was realized that supersymmetric theories can naturally solve the gauge hierarchy problem, explaining why the electroweak scale $M_{\rm EW} \sim 10^2$ GeV is so much smaller than the GUT or Planck scales. In this version of supersymmetry, which has now become the standard one, the gravitino mass will typically be $m_{3/2} \sim$ TeV; and the late decay of even a relatively small number of such massive particles can wreck BBN and/or the thermal spectrum of the CBR. The only way to prevent this is to make sure that the reheating temperature after inflation is sufficiently low: $T_{\rm RH} \lesssim 2 \times 10^9$ GeV (for $m_{3/2} =$ TeV) (Ellis, Kim, & Nanopoulos 1984, Ellis et al. 1992).

Interference between the tree-level (a) and one-loop (b) diagrams with complex Yukawa couplings can provide the requisite source of CP violation for GUT baryogenesis. In viable models, to avoid unwanted cancellations, one must often assume that the two scalars are different or go to higher loops (c)

2. Electroweak baryogenesis.

Below the electroweak scale of ~ 100 GeV, the **sphaleron** quantum tunneling process that violates B and L conservation (but preserves B - L) in the Standard Model is greatly suppressed, by ~ $\exp(-2\pi/\alpha_W) \sim 10^{-65}$. But at T ~ 100 GeV this process can occur. It can satisfy all three Sakharov conditions, but it cannot produce a large enough B and L. However, it can easily convert L into a mixture of B and L (Leptogenesis).

When one quantizes the Standard Model, one finds that the baryon number current is not exactly conserved, but rather satisfies

$$\partial_\mu j^\mu_B = \frac{3}{16\pi^2} F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} = \frac{3}{8\pi^2}\mathrm{Tr}F_{\mu\nu}\tilde{F}_{\mu\nu}$$

The same parity-violating term occurs in the divergence of the lepton number current, so the difference (the B - L current) is exactly conserved. The parity-violating term is a total divergence

$\mathrm{Tr}F_{\mu\nu}\tilde{F}_{\mu\nu} = \partial_\mu K^\mu$ where $K^\mu = \epsilon^{\mu\nu\rho\sigma} tr[F_{\nu\rho}A_\sigma + \frac{2}{3}A_\nu A_\rho A_\sigma]$, so

$\tilde{j} = j^\mu_B - \frac{3g^2}{8\pi^2}K^\mu$ is conserved. In perturbation theory (i.e. Feynman diagrams)

K^μ falls to zero rapidly at infinity, so B and L are conserved.

In abelian -- i.e. U(1) -- gauge theories, this is the end of the story. In non-abelian theories, however, there are non-perturbative field configurations, called instantons, which lead to violations of B and L. They correspond to calculation of a tunneling amplitude. To understand what the tunneling process is, one must consider more carefully the ground state of the field theory. Classically, the ground states are field configurations for which the energy vanishes. The trivial solution of this condition is A = 0, where A is the vector potential, which is the only possibility in U(1). But a "pure gauge" is also a solution, where

$$\vec{A} = \frac{1}{i} g^{-1} \vec{\nabla} g,$$

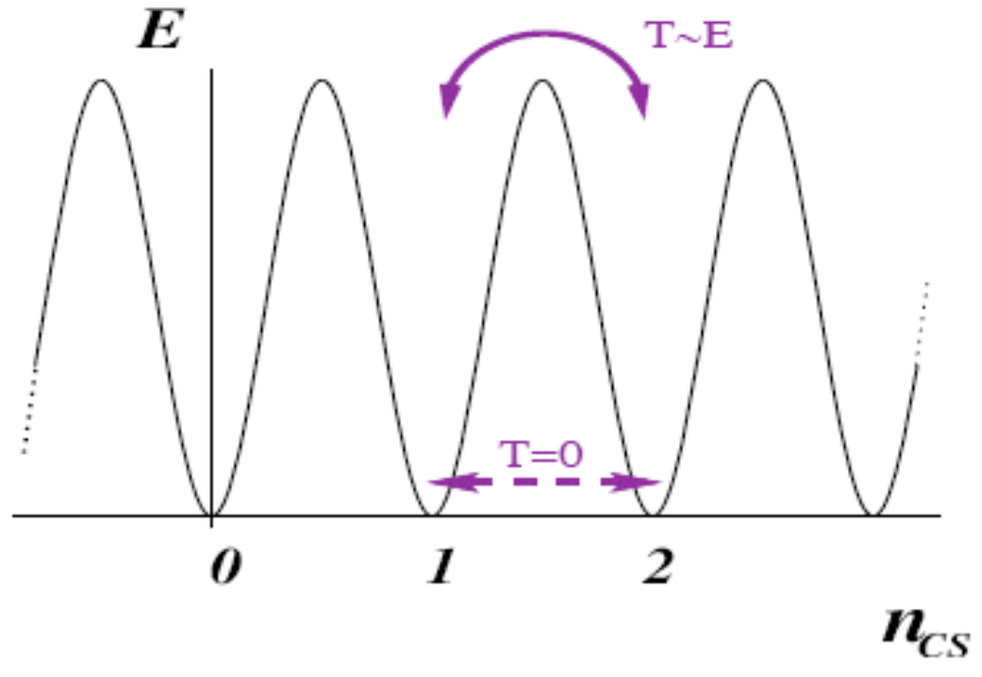
where g is a gauge transformation matrix. There is a class of gauge transformations g, labeled by a discrete index n, which must also be considered. These have the form

$$g_n(\vec{x}) = e^{in f(\vec{x}) \hat{x} \cdot \tau / 2} \text{ where } f(x) \to 2\pi \text{ as } \vec{x} \to \infty, \text{ and } f(\vec{x}) \to 0 \text{ as } \vec{x} \to 0.$$

The ground states are labeled by the index n. If we evaluate the integral of the current we obtaK^μ quantity known as the Chern-Simons number

$$n_{CS} = \frac{1}{16\pi^2} \int d^3x K^o = \frac{2/3}{16\pi^2} \int d^3x \epsilon_{ijk} Tr(g^{-1}\partial_i g g^{-1}\partial_j g g^{-1}\partial_k g). \text{ For } g = g_n, n_{CS} = n.$$

Schematic Yang-Mills vacuum structure. At zero temperature, the instanton transitions between vacua with different Chern-Simons numbers are suppressed. At finite temperature, these transitions can proceed via sphalerons.



In tunneling processes which change the Chern-Simons number, because of the anomaly, the baryon and lepton numbers will change. The exponential suppression found in the instanton calculation is typical of tunneling processes, and in fact the instanton calculation is nothing but a field-theoretic WKB calculation. The probability that a single proton has decayed through this process in the history of the universe is infinitesimal. But this picture suggests that, at finite temperature, the rate should be larger. One can determine the height of the barrier separating configurations of different n_{CS} by looking for the field configuration which corresponds to sitting on top of the barrier. This is a solution of the static equations of motion with finite energy. It is known as a "sphaleron". It follows that when the temperature is of order the ElectroWeak scale ~ 100 GeV, B and L violating (but B - L conserving) processes can proceed rapidly.

This result leads to three remarks:

1. If in the early universe, one creates baryon and lepton number, but no net $B - L$, B and L will subsequently be lost through sphaleron processes.

2. If one creates a net $B - L$ (e.g. creates a lepton number) the sphaleron process will leave both baryon and lepton numbers comparable to the original $B - L$. This realization is crucial to the idea of Leptogenesis.

3. The Standard Model satisfies, by itself, all of the conditions for baryogenesis. However, detailed calculations show that in the Standard Model the size of the baryon and lepton numbers produced are much too small to be relevant for cosmology, both because the Higgs boson is more massive than ~ 80 GeV and because the CKM CP violation is much too small. In supersymmetric extensions of the Standard Model it is possible that a large enough matter-antimatter asymmetry might be generated, but the parameter space for this is extremely small. (See Dine and Kusenko for details and references.)

This leaves Leptogenesis and Affleck-Dine baryogenesis as the two most promising possibilities. What is exciting about each of these is that, if they are operative, they have consequences for experiments which will be performed at accelerators over the next few years.

3. Leptogenesis.

There is now compelling experimental evidence that neutrinos have mass, both from solar and atmospheric neutrino experiments and accelerator and reactor experiments. The masses are tiny, fractions of an eV. The "see-saw mechanism" is a natural way to generate such masses. One supposes that in addition to the neutrinos of the Standard Model, there are some SU(2)xU(1)-singlet neutrinos, N. Nothing forbids these from obtaining a large mass. This could be of order M_{GUT}, for example, or a bit smaller. These neutrinos could also couple to the left handed doublets ν_L, just like right handed charged leptons. Assuming that these couplings are not particularly small, one would obtain a mass matrix, in the {N, ν_L} basis, of the form

$$M_\nu = \begin{pmatrix} M & M_W \\ M_W^T & 0 \end{pmatrix}$$

This matrix has an eigenvalue $\frac{M_W^2}{M}$.

The latter number is of the order needed to explain the light neutrino masses for M ~ 10^{13} GeV or so, i.e. not wildly different than the GUT scale and other scales which have been proposed for new physics. For **leptogenesis** (Fukugita and Yanagida, 1986), what is important in this model is that the couplings of N break lepton number. N is a heavy particle; it can decay both to h + v and h + v-bar, for example. The partial widths to each of these final states need not be the same. CP violation can enter through phases in the Yukawa couplings and mass matrices of the N's.

As the universe cools through temperatures of order the of masses of the N's, they drop out of equilibrium, and their decays can lead to an excess of neutrinos over antineutrinos. Detailed predictions can be obtained by integrating a suitable set of Boltzmann equations. These decays produce a net lepton number, but not baryon number (and hence a net B − L). The resulting lepton number will be further processed by sphaleron interactions, yielding a net lepton and baryon number (recall that sphaleron interactions preserve B − L, but violate B and L separately). Reasonable values of the neutrino parameters give asymmetries of the order we seek to explain.

It is interesting to ask: assuming that these processes are the source of the observed asymmetry, how many parameters which enter into the computation can be measured, i.e. can we relate the observed number to microphysics. It is likely that, over time, many of the parameters of the light neutrino mass matrices, including possible CP-violating effects, will be measured. But while these measurements determine some of the couplings and masses, they are not, in general, enough. In order to give a precise calculation, analogous to the calculations of nucleosynthesis, of the baryon number density, one needs additional information about the masses of the fields N. One either requires some other (currently unforseen) experimental access to this higher scale physics, or a compelling theory of neutrino mass in which symmetries, perhaps, reduce the number of parameters.

4. Production by coherent motion of scalar fields (the Affleck-Dine mechanism)

The formation of an AD condensate can occur quite generically in cosmological models. Also, the AD scenario potentially can give rise simultaneously to the ordinary matter and the dark matter in the universe. This can explain why the amounts of luminous and dark matter are surprisingly close to each other, within one order of magnitude. If the two entities formed in completely unrelated processes (for example, the baryon asymmetry from leptogenesis, while the dark matter from freeze-out of neutralinos), the observed relation $\Omega_{DARK} \sim \Omega_{baryon}$ is fortuitous.

In supersymmetric theories, the ordinary quarks and leptons are accompanied by scalar fields. These scalar fields carry baryon and lepton number. A coherent field, i.e., a large classical value of such a field, can in principle carry a large amount of baryon number. As we will see, it is quite plausible that such fields were excited in the early universe. To understand the basics of the mechanism, consider first a model with a single complex scalar field. Take the Lagrangian to be

$$\mathcal{L} = |\partial_\mu \phi|^2 - m^2 |\phi|^2$$

This Lagrangian has a symmetry, $\varphi \rightarrow e^{i\alpha} \varphi$, and a corresponding conserved current, which we will refer to as baryon current:

$$j_B^\mu = i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*).$$

It also possesses a “CP” symmetry: $\varphi \leftrightarrow \varphi*$. With supersymmetry in mind, we will think of m as of order M_W.

Let us add interactions in the following way, which will closely parallel what happens in the supersymmetric case. Include a set of quartic couplings:

$$\mathcal{L}_I = \lambda|\phi|^4 + \epsilon\phi^3\phi^* + \delta\phi^4 + c.c.$$

These interactions clearly violate B. For general complex ε and δ, they also violate CP. In supersymmetric theories, as we will shortly see, the couplings will be extremely small. In order that these tiny couplings lead to an appreciable baryon number, it is necessary that the fields, at some stage, were very large.

To see how the cosmic evolution of this system can lead to a non-zero baryon number, first note that at very early times, when the Hubble constant, $H \gg m$, the mass of the field is irrelevant. It is thus reasonable to suppose that at this early time $\varphi = \varphi_o \gg 0$. How does the field then evolve? First ignore the quartic interactions. In the expanding universe, the equation of motion for the field is as usual

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\partial V}{\partial \phi} = 0.$$

At very early times, $H \gg m$, and so the system is highly overdamped and essentially frozen at φ_o. At this point, B = 0.

Once the universe has aged enough that H ≪ m, φ begins to oscillate. Substituting H = ½ t or H = ⅔ t for the radiation and matter dominated eras, respectively, one finds that

$$\phi = \begin{cases} \frac{\phi_o}{(mt)^{3/2}} \sin(mt) & (\text{radiation}) \\ \frac{\phi_o}{(mt)} \sin(mt) & (\text{matter}). \end{cases}$$

In either case, the energy behaves, in terms of the scale factor, R(t), as

$$E \approx m^2 \phi_o^2 (\frac{R_o}{R})^3$$

Now let's consider the effects of the quartic couplings. Since the field amplitude damps with time, their significance will decrease with time. Suppose, initially, that φ = φ₀ is real. Then the imaginary part of φ satisfies, in the approximation that ε and δ are small,

$$\ddot{\phi}_i + 3H\dot{\phi}_i + m^2\phi_i \approx \mathrm{Im}(\epsilon + \delta)\phi_r^3.$$

For large times, the right hand falls as $t^{-9/2}$, whereas the left hand side falls off only as $t^{-3/2}$. As a result, baryon number violation becomes negligible. The equation goes over to the free equation, with a solution of the form

$$\phi_i = a_r \frac{\mathrm{Im}(\epsilon+\delta)\phi_o^3}{m^2(mt)^{3/4}} \sin(mt + \delta_r) \quad (\text{radiation}), \qquad \phi_i = a_m \frac{\mathrm{Im}(\epsilon+\delta)\phi_o^3}{m^3 t} \sin(mt + \delta_m) \quad (\text{matter}),$$

The constants can be obtained numerically, and are of order unity

$$a_r = 0.85 \qquad a_m = 0.85 \qquad \delta_r = -0.91 \qquad \delta_m = 1.54.$$

But now we have a non-zero baryon number; substituting in the expression for the current,

$$n_B = 2a_r \mathrm{Im}(\epsilon + \delta) \frac{\phi_o^2}{m(mt)^2} \sin(\delta_r + \pi/8) \quad \text{(radiation)}$$

$$n_B = 2a_m \mathrm{Im}(\epsilon + \delta) \frac{\phi_o^2}{m(mt)^2} \sin(\delta_m) \quad \text{(matter)}.$$

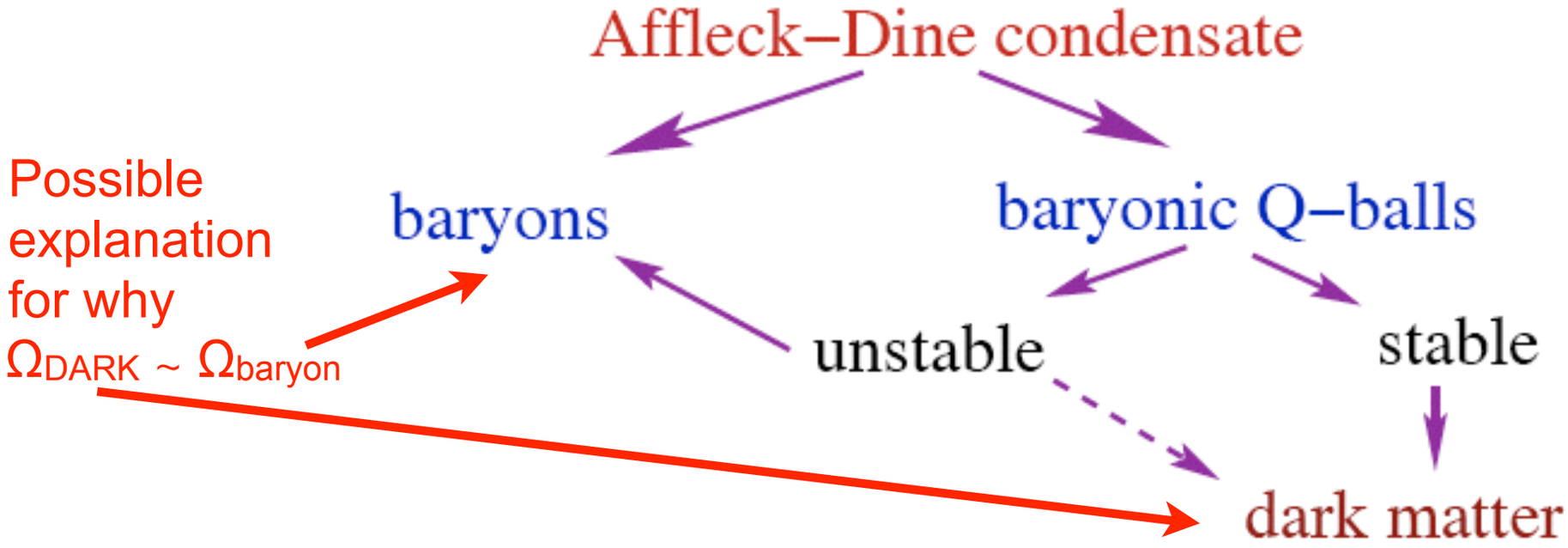
Two features of these results should be noted. First, if ε and δ vanish, n_B vanishes. If they are real, and φ_o is real, n_B vanishes. It is remarkable that the Lagrangian parameters can be real, and yet φ_o can be complex, still giving rise to a net baryon number. Supersymmetry breaking in the early universe can naturally lead to a very large value for a scalar field carrying B or L. Finally, as expected, n_B is conserved at late times.

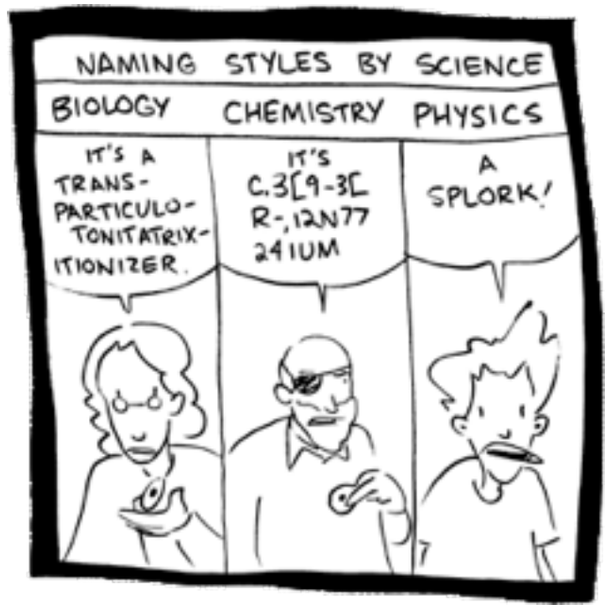
This mechanism for generating baryon number could be considered without supersymmetry. In that case, it begs several questions:

- What are the scalar fields carrying baryon number?
- Why are the φ^4 terms so small?
- How are the scalars in the condensate converted to more familiar particles?

In the context of supersymmetry, there is a natural answer to each of these questions. First, there are scalar fields (squarks and sleptons) carrying baryon and lepton number. Second, in the limit that supersymmetry is unbroken, there are typically directions in the field space in which the quartic terms in the potential vanish. Finally, the scalar quarks and leptons will be able to decay (in a baryon and lepton number conserving fashion) to ordinary quarks.

In addition to topologically stable solutions to the field equations such as strings or monopoles, it is sometimes also possible to find non-topological solutions, called Q-balls, which can form as part of the Affleck-Dine condensate. These are usually unstable and could decay to the dark matter, but in some theories they are stable and could be the dark matter. The various possibilities are summarized as follows:





The parameter space of the MSSM consistent with LSP dark matter is very different, depending on whether the LSPs froze out of equilibrium or were produced from the evaporation of AD baryonic Q-balls. If supersymmetry is discovered, one will be able to determine the properties of the LSP experimentally. This will, in turn, provide some information on how the dark-matter SUSY particles could be produced. The discovery of a Higgsino-like LSP would be evidence in favor of Affleck–Dine baryogenesis. This is a way in which we might be able to establish the origin of matter-antimatter asymmetry.

Review of mechanisms that have been proposed to generate the baryon asymmetry:

1. **GUT Baryogenesis**. Grand Unified Theories unify the gauge interactions of the strong, weak and electromagnetic interactions in a single gauge group. They inevitably violate baryon number, and they have heavy particles, with mass of order $M_{GUT} \approx 10^{16}$ GeV, whose decays can provide a departure from equilibrium. The main objections to this possibility come from issues associated with inflation. While there does not exist a compelling microphysical model for inflation, in most models, the temperature of the universe after reheating is well below M_{GUT}. But even if it were very large, there would be another problem. Successful unification requires supersymmetry, which implies that the graviton has a spin-3/2 partner, called the gravitino. In most models for supersymmetry breaking, these particles have masses of order TeV, and are very long lived. Even though these particles are weakly interacting, too many gravitinos are produced unless the reheating temperature is well below the unification scale -- too low for GUT baryogenesis to occur.

2. **Electroweak baryogenesis**. The Standard Model satisfies all of the conditions for baryogenesis, but any baryon asymmetry produced is far too small to account for observations. In certain extensions of the Standard Model, it is possible to obtain an adequate asymmetry, but in most cases the allowed region of parameter space is very small.

3. **Leptogenesis**. The possibility that the weak interactions will convert some lepton number to baryon number means that if one produces a large lepton number at some stage, this will be processed into a net baryon and lepton number at the electroweak phase transition. The observation of neutrino masses makes this idea highly plausible. Many but not all of the relevant parameters can be directly measured.

4. **Production by coherent motion of scalar fields (the Affleck-Dine mechanism)**, which can be highly efficient, might well be operative if nature is supersymmetric.