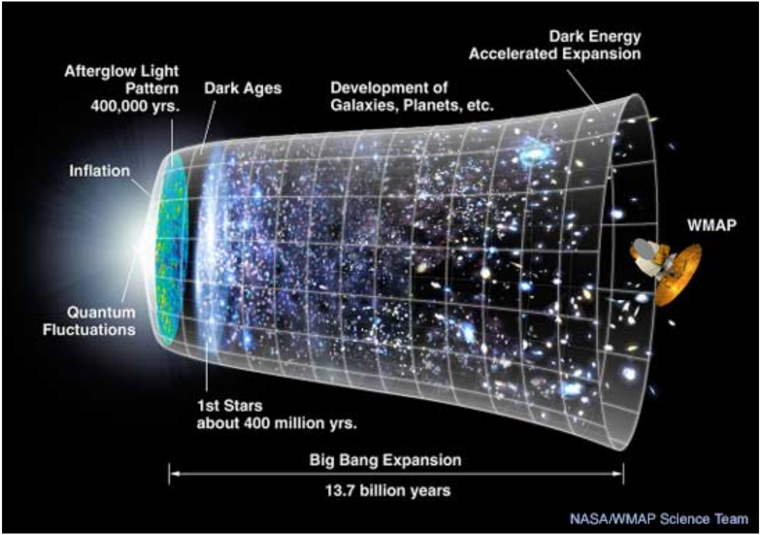
The simplest viable inflationary models

Alexei A. Starobinsky

Landau Institute for Theoretical Physics RAS,
Moscow - Chernogolovka, Russia

XVI-DIAS-TH Winter School "Actual Cosmology"
JINR, Dubna, Russia, 02-03.02.2021


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
400,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
WMAP
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years
NASA/WMAP Science Team

Four epochs of the history of the Universe

$H \equiv \frac{\dot{a}}{a}$ where $a(t)$ is a scale factor of an isotropic homogeneous spatially flat universe (a Friedmann-Lemaître-Robertson-Walker background):

$$ds^2 = dt^2 - a^2(t)(dx^2 + dy^2 + dz^2) + small\ perturbations$$

The history of the Universe in one line: four main epochs

$$? \longrightarrow DS \Longrightarrow FLRWRD \Longrightarrow FLRWMD \Longrightarrow \overline{DS} \longrightarrow ?$$

Geometry

$$|\dot{H}| << H^2 \Longrightarrow H = \frac{1}{2t} \Longrightarrow H = \frac{2}{3t} \Longrightarrow |\dot{H}| << H^2$$

Physics

$$p \approx -\rho \Longrightarrow p = \rho/3 \Longrightarrow p \ll \rho \Longrightarrow p \approx -\rho$$

Duration in terms of the number of e-folds $\ln(a_{fin}/a_{in})$

> 60 ~ 55 7.5 0.5

Principal epochs of the Universe evolution – before 1979

The history of the Universe in one line: two principal epochs

$$? \longrightarrow FLRWRD \Longrightarrow FLRWMD \longrightarrow ?$$

Geometry

$$H = \frac{1}{2t} \Longrightarrow H = \frac{2}{3t}$$

Physics

$$p = \rho/3 \Longrightarrow p \ll \rho$$

Inflation

The inflationary hypothesis (I follow the way how it was introduced in my JETP Lett.1979 and PLB 1980 papers):

Some part of the world which includes all its presently observable part was as much symmetric as possible during some period in the past - both with respect to the geometrical background and to the state of all quantum fields (no particles).

Non-universal (due to the specific initial condition) explanation of the cosmological arrow to time - entropy (in some not well defined sense) can only grow after inflation.

Still this state is an intermediate attractor for a set of pre-inflationary initial conditions with a non-zero measure. Also it is not a unique one, there exists a class of such states leading to the same observable predictions.

Successive realization of this idea is based on the two more detailed and independent assumptions.

1. Existence of a metastable quasi-de Sitter stage in our remote part which preceded the hot Big Bang. During it, the expansion of the Universe was accelerated and close to the exponential one, $|\dot{H}| \ll H^2$.

2. The origin of all inhomogeneities in the present Universe is the effect of gravitational creation of pairs of particles - antiparticles and field fluctuations during inflation from the adiabatic vacuum (no-particle) state for Fourier modes covering all observable range of scales (and possibly somewhat beyond).

NB. This effect is the same as particle creation by black holes, but no problems with the loss of information, 'firewalls', trans-Planckian energy etc. in cosmology, as far as observational predictions are calculated.

Remark regarding these initial conditions for perturbations: they are *not* in the Bunch-Davies state in the rigorous sense, since they may not be imposed for arbitrary large scales. As a consequence, inflationary models typically does *not* predict regular behaviour at spatial infinity both during and after inflation ("multiverse").

Existing analogies in other areas of physics.

1. The present dark energy, though the required degree of metastability for the primordial dark energy is much more than is proved for the present one (more than 60 e-folds vs. ~ 3).
2. Creation of electrons and positrons in an external electric field.

Outcome of inflation

In the super-Hubble regime ($k \ll aH$) in the coordinate representation:

$$ds^2 = dt^2 - a^2(t)(\delta_{lm} + h_{lm})dx^l dx^m, \quad l, m = 1, 2, 3$$

$$h_{lm} = 2\mathcal{R}(\mathbf{r})\delta_{lm} + \sum_{a=1}^{2} g^{(a)}(\mathbf{r})\, e_{lm}^{(a)}$$

$$e_l^{l(a)} = 0, \;\; g^{(a)}_{,l}\, e_m^{l(a)} = 0, \;\; e_{lm}^{(a)}\, e^{lm(a)} = 1$$

$\mathcal{R}$ describes primordial scalar perturbations, g – primordial tensor perturbations (primordial gravitational waves (GW)).

The most important quantities:

$$n_s(k) - 1 \equiv \frac{d \ln P_{\mathcal{R}}(k)}{d \ln k}, \quad r(k) \equiv \frac{P_g}{P_{\mathcal{R}}}$$

Existence of constant modes

For FLRW models filled by ideal fluids, it was known already to Lifshitz (1946). For a wide class of modified scalar-tensor gravity theories, it was proved in A. A. Starobinsky, S. Tsujikawa and J. Yokoyama, Nucl. Phys. B 610, 383 (2001). However, their existence is much more general. From the mathematical point of view, constant modes appear simply due to the existence of non-degenerate solutions of the same gravity models in the isotropic and spatially flat FLRW space-time. By construction, these solutions always have 3 non-physical (gauge) arbitrary constants of integration due to the possibility of arbitrary and independent rescaling of all spatial coordinates. Making these constants slightly inhomogeneous converts them to the leading terms of physical constant modes (one scalar and two tensor ones). Moreover, it straightforwardly follows from this that these constants (now functions of spatial coordinates) need not be small, they can be arbitrarily large.

In fact, metric perturbations h_{lm} are quantum (operators in the Heisenberg representation) and remain quantum up to the present time. But, after omitting of a very small part, decaying with time, they become commuting and, thus, equivalent to classical (c-number) stochastic quantities with the Gaussian statistics (up to small terms quadratic in $\mathcal{R}$, g).

In particular:

$$\hat{\mathcal{R}}_k = \mathcal{R}_k \, i(\hat{a}_{\mathbf{k}} - \hat{a}_{\mathbf{k}}^\dagger) + \mathcal{O}\left((\hat{a}_{\mathbf{k}} - \hat{a}_{\mathbf{k}}^\dagger)^2\right) + ... + \mathcal{O}(10^{-100})(\hat{a}_{\mathbf{k}} + \hat{a}_{\mathbf{k}}^\dagger) + ,,,$$

The last term is time dependent, it is affected by physical decoherence and may become larger, but not as large as the second term.

Remaining quantum coherence: deterministic correlation between $\mathbf{k}$ and $-\mathbf{k}$ modes - shows itself in the appearance of acoustic oscillations (primordial oscillations in case of GW).

All these predictions are beyond semiclassical gravity!

Semiclassical gravity: space-time metric g_{ik} is not quantized and

$$\frac{1}{8\pi G}\left(R^{\nu}_{\mu} - \frac{1}{2}\,\delta^{\nu}_{\mu} R\right)(g_{ik}) = \left\langle \hat{T}^{\nu}_{\mu} \right\rangle$$

Instead,

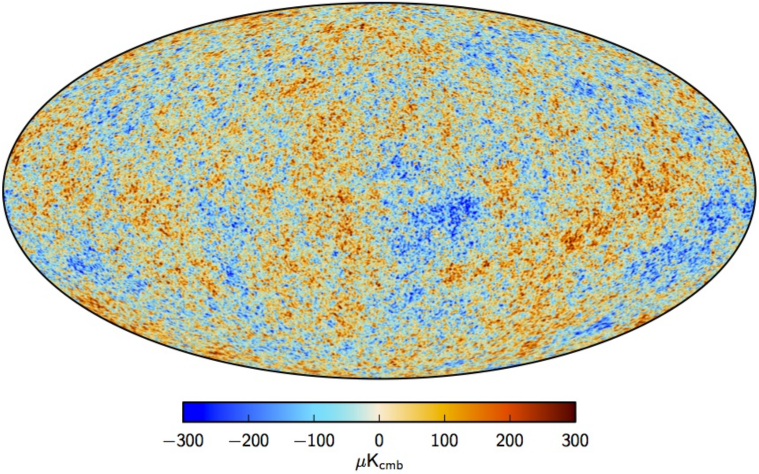
$$\frac{1}{8\pi G}\left(\hat{R}^{\nu}_{\mu} - \frac{1}{2}\,\delta^{\nu}_{\mu} \hat{R}\right)(\hat{g}_{ik}) = \hat{T}^{\nu}_{\mu}$$

is used.

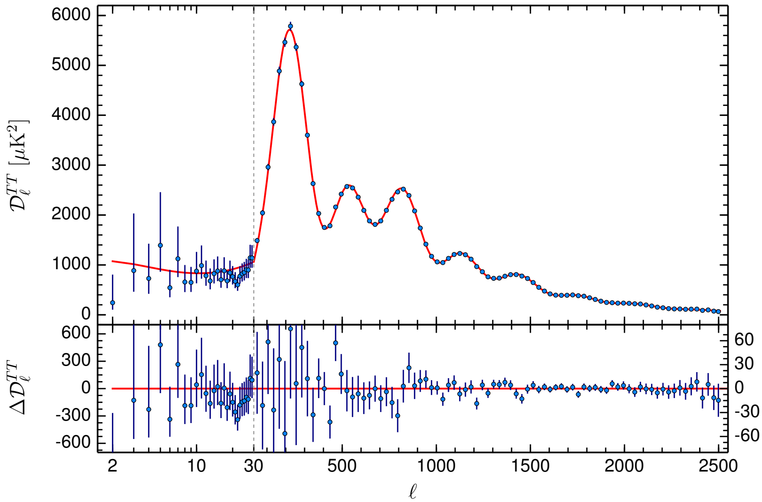
$\langle \mathcal{R} \rangle = 0$ does not mean the absence of perturbations.

CMB temperature anisotropy

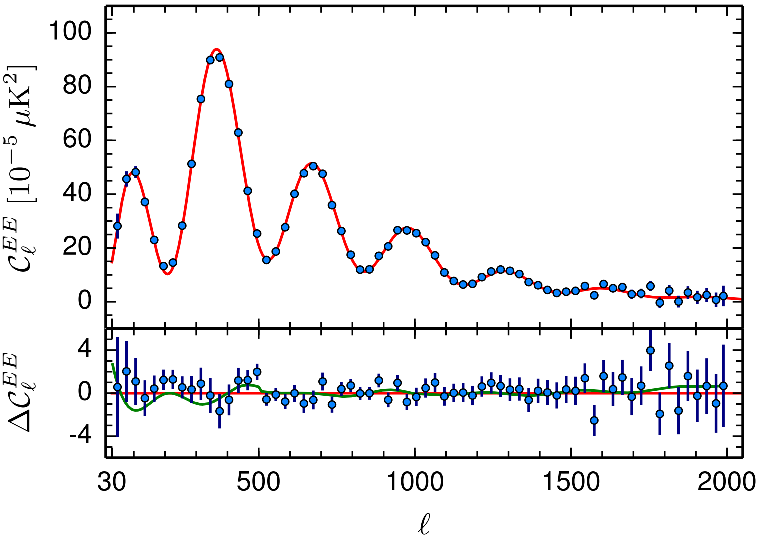
Planck-2015: P. A. R. Ade et al., arXiv:1502.01589



CMB temperature anisotropy multipoles



CMB E-mode polarization multipoles



New cosmological parameters relevant to inflation

Now we have numbers: N. Agranim et al., arXiv:1807.06209

The primordial spectrum of scalar perturbations has been measured and its deviation from the flat spectrum $n_s = 1$ in the first order in $|n_s - 1| \sim N_H^{-1}$ has been discovered (using the multipole range $\ell > 40$):

$$< \mathcal{R}^2(\mathbf{r}) >= \int \frac{P_{\mathcal{R}}(k)}{k}\, dk, \;\; P_{\mathcal{R}}(k) = (2.10 \pm 0.03)\cdot 10^{-9} \left(\frac{k}{k_0}\right)^{n_s - 1}$$

$$k_0 = 0.05 \ \mathrm{Mpc}^{-1}, \quad n_s - 1 = -0.035 \pm 0.004$$

Two fundamental observational constants of cosmology in addition to the three known ones (baryon-to-photon ratio, baryon-to-matter density and the cosmological constant). Existing inflationary models can predict (and predicted, in fact) one of them, namely $n_s - 1$, relating it finally to $N_H = \ln \frac{k_B T_\gamma}{\hbar H_0} \approx 67.2$. (note that $(1 - n_s)N_H \sim 2$).

Physical scales related to inflation

"Naive" estimate where I use the reduced Planck mass $\tilde{M}_{Pl} = (8\pi G)^{-1}$.

I. Curvature scale

$$H \sim \sqrt{P_{\mathcal{R}}}\tilde{M}_{Pl} \sim 10^{14}\mathrm{GeV}$$

II. Inflaton mass scale

$$|m_{infl}| \sim H\sqrt{|1 - n_s|} \sim 10^{13}\mathrm{GeV}$$

New range of mass scales significantly less than the GUT scale.

Inflationary slow-roll dynamics

The crucial element of all existing viable inflationary models: slow-roll evolution of the Hubble factor and scalar fields during inflation:

$$|\ddot{\phi}| \ll H|\dot{\phi}|, \quad \dot{\phi}^2 \ll V(\phi), \quad |\dot{H}| \ll H^2$$

.

Necessary conditions for scalar field models in GR:
$|V'| \ll \kappa V, \quad |V''| \ll \kappa^2 V$. Then

$$H^2 \approx \frac{\kappa^2 V}{3}, \quad \dot{\phi} \approx -\frac{V'}{3H}, \quad N \equiv \ln \frac{a_f}{a} \approx \kappa^2 \int_{\phi_f}^{\phi} \frac{V}{V'} \, d\phi$$

First obtained in A. A. Starobinsky, Sov. Astron. Lett. 4, 82 (1978) in the $V = \frac{m^2\phi^2}{2}$ case and for a bouncing model.

Kinematic origin of scalar perturbations

Local duration of inflation in terms of $N_{tot} = \ln\left(\frac{a(t_{fin})}{a(t_{in})}\right)$ is different in different points of space: $N_{tot} = N_{tot}(\mathbf{r})$. Then

$$\mathcal{R}(\mathbf{r}) = \delta N_{tot}(\mathbf{r})$$

Correct generalization to the non-linear case: the space-time metric after the end of inflation at super-Hubble scales

$$ds^2 = dt^2 - a^2(t)e^{2N_{tot}(\mathbf{r})}(dx^2 + dy^2 + dz^2)$$

First derived in A. A. Starobinsky, Phys. Lett. B 117, 175 (1982) in the case of one-field inflation.

Visualizing small differences in the number of e-folds

Duration of inflation in terms of e-folds was finite for all points inside our past light cone. For $\ell \lesssim 50$, neglecting the Silk and Doppler effects, as well as the ISW effect due the presence of dark energy,

$$\frac{\Delta T(\theta,\phi)}{T_\gamma} = -\frac{1}{5}\mathcal{R}(r_{LSS},\theta,\phi) = -\frac{1}{5}\delta N_{tot}(r_{LSS},\theta,\phi)$$

For $n_s = 1$,

$$\ell(\ell+1)C_{\ell,s} = \frac{2\pi}{25}P_\zeta$$

For $\frac{\Delta T}{T} \sim 10^{-5}$, $\delta N \sim 5\times10^{-5}$, and for $H \sim 10^{14}\,\mathrm{GeV}$, $\delta t \sim 5 t_{Pl}$!

Planck time intervals are seen by the naked eye!

Three general directions in criticizing inflation

1. "Inflation is not falsifiable and cannot predict anything". Reply. Wrong. Any concrete inflationary model is falsifiable, and the most of proposed models have been already falsified. The less is the number of free parameters (to be fixed by observations only) of remaining viable models, the larger is their predictive power. The most predictive models have only one free parameter at present.

2. "Beginning of (isotropic and homogeneous) inflation requires extreme fine-tuning of pre-inflationary evolution". Reply. Wrong. Intermediate inflationary attractor is generic and inhomogeneous, it approaches (quasi-)de Sitter only locally, inside the Hubble radius. As a result, like black holes in GR, inflationary patches always form somewhere.

3. "Eternal inflation leads to multiverse in which no quantitative predictions can be made".
Reply. Not relevant (if not wrong). Irrespective of some problems with making predictions in this case (which arise due to unjustified manipulations with conditional probabilities to some extent), eternal inflation is an unfinished, still evolving process occurring outside our past (and future) light cone. Thus, it does not affect what we see in our Universe up to fantastically (exponentially) small terms which are always neglected in all quantitative considerations of eternal inflation.

Here the term "multiverse" is used in the narrow sense: the matter action S_m is the same for all post-inflationary universes.

The simplest models producing the observed scalar slope

1. The $R+R^2$ model (Starobinsky 1980):

$$\mathcal{L}=\frac{f(R)}{16\pi G},\quad f(R)=R+\frac{R^2}{6M^2},\quad M_{\rm Pl}^2=G^{-1}$$

$$M=2.6\times10^{-6}\left(\frac{55}{N}\right)M_{\rm Pl}\approx3.1\times10^{13}\,\mathrm{GeV}$$

$$n_s-1=-\frac{2}{N}\approx-0.036,\quad r=\frac{12}{N^2}\approx0.004,\quad n_t=-\frac{r}{8}$$

$$N=\ln\frac{k_f}{k}=\ln\frac{T_\gamma}{k}-\mathcal{O}(10),\quad H_{dS}(N=55)=1.4\times10^{14}\,\mathrm{GeV}$$

2. The same prediction from a scalar field model with $V(\phi)=\frac{\lambda\phi^4}{4}$ at large ϕ and strong non-minimal coupling to gravity $\xi R\phi^2$ with $\xi<0,\ |\xi|\gg1$ (Spokoiny 1984), including the Higgs inflationary model (Bezrukov and Shaposhnikov 2008).

The simplest purely geometrical inflationary model

$$\mathcal{L} = \frac{R}{16\pi G} + \frac{N^2}{288\pi^2 P_{\mathcal{R}}(k)} R^2 + (\text{small rad. corr.})$$

$$= \frac{R}{16\pi G} + 5.1 \times 10^8 \, R^2 + (\text{small rad. corr.})$$

The quantum effect of creation of particles and field fluctuations works twice in this model:
a) at super-Hubble scales during inflation, to generate space-time metric fluctuations;
b) at small scales after inflation, to provide scalaron decay into pairs of matter particles and antiparticles (AS, 1980, 1981).

Weak dependence of the time t_r when the radiation dominated stage begins:

$$N(k) \approx N_H + \ln \frac{a_0 H_0}{k} - \frac{1}{3} \ln \frac{M_{\rm Pl}}{M} - \frac{1}{6} \ln(M_{\rm Pl} t_r)$$

Evolution of the $R + R^2$ model

1. During inflation ($H \gg M$):

$$H = \frac{M^2}{6}(t_f - t) + \frac{1}{6(t_f - t)} + ..., \quad |\dot{H}| \ll H^2$$

(for the derivation of the second term in the rhs - see A. S. Koshelev et al., JHEP 1611 (2016) 067).

2. After inflation ($H \ll M$):

$$a(t) \propto t^{2/3}\left(1 + \frac{2}{3Mt}\sin M(t - t_1)\right)$$

The most effective decay channel: into minimally coupled scalars and the longitudinal mode of vector bosons with $m \ll M$. In the first case the formula

$$\frac{1}{\sqrt{-g}}\frac{d}{dt}(\sqrt{-g}\,n_s) = \frac{R^2}{576\pi}$$

(Ya. B. Zeldovich and A. A. Starobinsky, JETP Lett. 26, 252 (1977)) can be used for simplicity, but the full integral-differential system of equations for the Bogoliubov α_k, β_k coefficients and the average EMT was in fact solved in AS (1981). Scalaron decay into graviton pairs is suppressed (A. A. Starobinsky, JETP Lett. 34, 438 (1981)).

For this channel of the scalaron decay:

$$\Gamma = \frac{M^3}{24 M_{\rm Pl}^2}, \quad N(k) \approx N_H + \ln\frac{a_0 H_0}{k} - \frac{5}{6}\ln\frac{M_{\rm Pl}}{M}$$

that gives $N(k = 0.002\,\mathrm{Mpc}^{-1}) \approx 54$. For the Higgs and the mixed R^2-Higgs models, $N(k = 0.002\,\mathrm{Mpc}^{-1}) \approx 58$, the increase is mainly due to the large Higgs non-minimal coupling.

Possible microscopic origins of this phenomenological model.

1. Follow the purely geometrical approach and consider it as the specific case of the fourth order gravity in 4D

$$\mathcal{L} = \frac{R}{16\pi G} + AR^2 + BC_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} + (\text{small rad. corr.})$$

for which $A \gg 1, \ A \gg |B|$. Approximate scale (dilaton) invariance and absence of ghosts in the curvature regime $A^{-2} \ll (RR)/M_P^4 \ll B^{-2}$.

One-loop quantum-gravitational corrections are small (their imaginary parts are just the predicted spectra of scalar and tensor perturbations), non-local and qualitatively have the same structure modulo logarithmic dependence on curvature.

2. Another, completely different way:

consider the $R+R^2$ model as an approximate description of GR + a non-minimally coupled scalar field with a large negative coupling ξ ($\xi_{conf} = \frac{1}{6}$) in the gravity sector:

$$\mathcal{L} = \frac{R}{16\pi G} - \frac{\xi R\phi^2}{2} + \frac{1}{2}\phi_{,\mu}\phi^{,\mu} - V(\phi), \quad \xi < 0, \quad |\xi| \gg 1 \ .$$

Geometrization of the scalar:

for a generic family of solutions during inflation, the scalar kinetic term can be neglected, so

$$\xi R\phi = -V'(\phi) + \mathcal{O}(|\xi|^{-1}) \ .$$

No conformal transformation, we remain in the the physical (Jordan) frame!

These solutions are the same as for $f(R)$ gravity with

$$\mathcal{L} = \frac{f(R)}{16\pi G}, \quad f(R) = R - \frac{\xi R \phi^2(R)}{2} - V(\phi(R)).$$

For $V(\phi) = \frac{\lambda(\phi^2 - \phi_0^2)^2}{4}$, this just produces $f(R) = \frac{1}{16\pi G}\left(R + \frac{R^2}{6M^2}\right)$ with $M^2 = \lambda/24\pi\xi^2 G$ and $\phi^2 = |\xi|R/\lambda$.

The same theorem is valid for a multi-component scalar field.

More generally, R^2 inflation (with an arbitrary n_s, r) serves as an intermediate dynamical attractor for a large class of scalar-tensor gravity models.

Inflation in $f(R)$ gravity

The simplest model of modified gravity (= geometrical dark energy) considered as a phenomenological macroscopic theory in the fully non-linear regime and non-perturbative regime.

$$S = \frac{1}{16\pi G} \int f(R)\sqrt{-g}\, d^4x + S_m$$

$$f(R) = R + F(R), \quad R \equiv R_\mu^\mu$$

Here $f''(R)$ is not identically zero. Usual matter described by the action S_m is minimally coupled to gravity.

Vacuum one-loop corrections depending on R only (not on its derivatives) are assumed to be included into $f(R)$. The normalization point: at laboratory values of R where the scalaron mass (see below) $m_s \approx \mathrm{const}$.

Metric variation is assumed everywhere. Palatini variation leads to a different theory with a different number of degrees of freedom.

Field equations

$$\frac{1}{8\pi G}\left(R_\mu^\nu - \frac{1}{2}\,\delta_\mu^\nu R\right) = -\left(T_{\mu\,(vis)}^\nu + T_{\mu\,(DM)}^\nu + T_{\mu\,(DE)}^\nu\right)\ ,$$

where $G = G_0 = const$ is the Newton gravitational constant measured in laboratory and the effective energy-momentum tensor of DE is

$$8\pi G T_{\mu\,(DE)}^\nu = F'(R)\,R_\mu^\nu - \frac{1}{2}\,F(R)\delta_\mu^\nu + \left(\nabla_\mu\nabla^\nu - \delta_\mu^\nu\nabla_\gamma\nabla^\gamma\right)F'(R)\,.$$

Because of the need to describe DE, de Sitter solutions in the absence of matter are of special interest. They are given by the roots $R = R_{dS}$ of the algebraic equation

$$Rf'(R) = 2f(R)\ .$$

The special role of $f(R) \propto R^2$ gravity: admits de Sitter solutions with any curvature.

Degrees of freedom

I. In quantum language: particle content.
1. Graviton – spin 2, massless, transverse traceless.
2. Scalaron – spin 0, massive, mass - R-dependent:
$m_s^2(R) = \frac{1}{3f''(R)}$ in the WKB-regime.

II. Equivalently, in classical language: number of free functions of spatial coordinates at an initial Cauchy hypersurface.
Six, instead of four for GR – two additional functions describe massive scalar waves.

Thus, $f(R)$ gravity is a non-perturbative generalization of GR.
It is equivalent to scalar-tensor gravity with $\omega_{BD} = 0$ if $f''(R) \neq 0$.

Transformation to the Einstein frame and back

In the Einstein frame, free particles of usual matter do not follow geodesics and atomic clocks do not measure proper time.

From the Jordan (physical) frame to the Einstein one:

$$g^E_{\mu\nu} = f' g^J_{\mu\nu}, \quad \kappa\phi = \sqrt{\frac{3}{2}} \ln f', \quad V(\phi) = \frac{f'R - f}{2\kappa^2 f'^2}$$

where $\kappa^2 = 8\pi G$.

Inverse transformation:

$$R = \left(\sqrt{6}\kappa \frac{dV(\phi)}{d\phi} + 4\kappa^2 V(\phi)\right) \exp\left(\sqrt{\frac{2}{3}}\kappa\phi\right)$$

$$f(R) = \left(\sqrt{6}\kappa \frac{dV(\phi)}{d\phi} + 2\kappa^2 V(\phi)\right) \exp\left(2\sqrt{\frac{2}{3}}\kappa\phi\right)$$

$V(\phi)$ should be at least C^1.

Background FRW equations in $f(R)$ gravity

$$ds^2 = dt^2 - a^2(t)\left(dx^2 + dy^2 + dz^2\right)$$

$$H \equiv \frac{\dot{a}}{a}\,, \quad R = 6(\dot{H} + 2H^2)$$

The trace equation (4th order)

$$\frac{3}{a^3}\frac{d}{dt}\left(a^3\frac{df'(R)}{dt}\right) - Rf'(R) + 2f(R) = 8\pi G(\rho_m - 3p_m)$$

The 0-0 equation (3d order)

$$3H\frac{df'(R)}{dt} - 3(\dot{H} + H^2)f'(R) + \frac{f(R)}{2} = 8\pi G\rho_m$$

Reduction to the first order equation

In the absence of spatial curvature and $\rho_m = 0$, it is always possible to reduce these equations to a first order one using either the transformation to the Einstein frame and the Hamilton-Jacobi-like equation for a minimally coupled scalar field in a spatially flat FLRW metric, or by directly transforming the 0-0 equation to the equation for $R(H)$:

$$\frac{dR}{dH} = \frac{(R - 6H^2)f'(R) - f(R)}{H(R - 12H^2)f''(R)}$$

See, e.g. H. Motohashi amd A. A. Starobinsky, Eur. Phys. J. C **77**, 538 (2017), but in the special case of the $R + R^2$ gravity this was found and used already in the original AS (1980) paper.

Analogues of large-field (chaotic) inflation: $F(R) \approx R^2 A(R)$ for $R \to \infty$ with $A(R)$ being a slowly varying function of R, namely

$$|A'(R)| \ll \frac{A(R)}{R}\ ,\ |A''(R)| \ll \frac{A(R)}{R^2}\ .$$

Analogues of small-field (new) inflation, $R \approx R_1$:

$$F'(R_1) = \frac{2F(R_1)}{R_1}\ ,\ F''(R_1) \approx \frac{2F(R_1)}{R_1^2}\ .$$

Thus, all inflationary models in $f(R)$ gravity are close to the simplest one over some range of R.

Perturbation spectra in slow-roll $f(R)$ inflationary models

Let $f(R) = R^2 A(R)$. In the slow-roll approximation $|\ddot{R}| \ll H|\dot{R}|$:

$$P_{\mathcal{R}}(k) = \frac{\kappa^2 A_k}{64\pi^2 A_k'^2 R_k^2}, \quad P_g(k) = \frac{\kappa^2}{12 A_k \pi^2}, \quad \kappa^2 = 8\pi G$$

$$N(k) = -\frac{3}{2}\int_{R_f}^{R_k} dR\, \frac{A}{A' R^2}$$

where the index k means that the quantity is taken at the moment $t = t_k$ of the Hubble radius crossing during inflation for each spatial Fourier mode $k = a(t_k)H(t_k)$.

Inflation in the mixed R^2-Higgs model

M. He, A. A. Starobinsky and J. Yokoyama, JCAP 1805, 064 (2018).

$$\mathcal{L} = \frac{1}{16\pi G}\left(R + \frac{R^2}{6M^2}\right) - \frac{\xi R \chi^2}{2} + \frac{1}{2}\chi_{,\mu}\chi^{,\mu} - \frac{\lambda\chi^4}{4}, \;\; \xi < 0, \;\; |\xi| \gg 1$$

Can be conformally transformed to GR with two interacting scalar fields in the Einstein frame. The effective two scalar field potential for the dual model:

$$U = e^{-2\alpha\phi}\left(\frac{\lambda}{4}\chi^4 + \frac{M^2}{2\alpha^2}\left(e^{\alpha\phi} - 1 + \xi\kappa^2\chi^2\right)^2\right)$$

$$\alpha = \sqrt{\frac{2}{3}}\,\kappa, \quad R = 3M^2\left(e^{\alpha\phi} - 1 + \xi\kappa^2\chi^2\right)$$

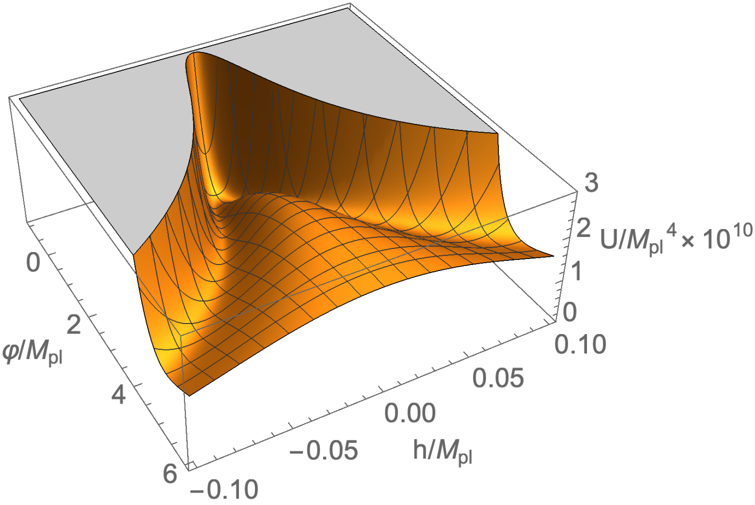
Attractiveness of the model from the quantum field theory point of view

The mixed R^2-Higgs model helps to remove some UV problems of the Higgs inflationary model and may be considered as its UV-completion up to the Planck energy if

$$\sqrt{\lambda} \lesssim \frac{|\xi| M}{M_{\rm Pl}} \lesssim 1$$

(see D. Gorbunov and A. Tokareva, Phys. Lett. B 788, 37 (2019)).

Effective potential in the Einstein frame



One-field inflation in the attractor regime

In the attractor regime during inflation:

$$\alpha\phi \gg 1, \quad \chi^2 \approx \frac{|\xi| R}{\lambda}, \quad e^{\alpha\phi} \approx \chi^2 \left(|\xi|\kappa^2 + \frac{\lambda}{3|\xi| M^2} \right)$$

that directly follows from the geometrization of the Higgs boson in the physical (Jordan) frame. Thus, we return to the $f(R) = R + \frac{R^2}{6M^2}$ model with the renormalized scalaron mass $M \to \tilde{M}$:

$$\frac{1}{\tilde{M}^2} = \frac{1}{M^2} + \frac{3\xi^2\kappa^2}{\lambda}$$

Double-field inflation reduces to the single $(R + R^2)$ one for the most of trajectories in the phase space.

Post-inflationary heating in the mixed R^2-Higgs model through particle creation

The most effective channel of reheating though particle creation: creation of longitudinal quanta of vector bosons with $m \ll \min(M, \sqrt{\lambda} M_{\mathrm{Pl}}/|\xi|)$. More effective than in the pure R^2 model, but less effective than in the pure Higgs case.

The simplified variant - creation of NG (phase direction) quanta of a complex Higgs-like scalar field: M. He, R. Jinno, K. Kamada, S. C. Park, A. A. Starobinsky and J. Yokoyama, Phys. Lett. B 791, 36 (2019) [arXiv:1812.10099]. Inflaton decay is not instant and occurs after a large number of scalaron oscillations. However, the Higgs field introduces anharmonic effects in these oscillations.

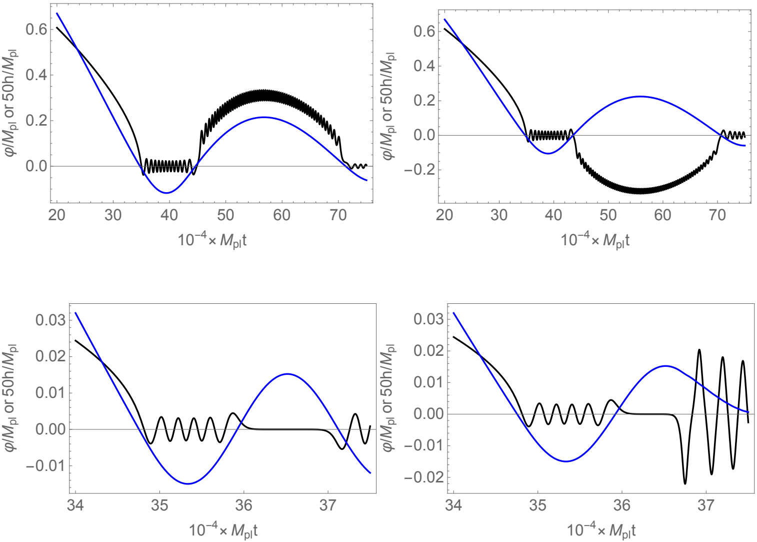
Post-inflationary heating in the mixed R^2-Higgs model through tachyonic preheating

M. He, R. Jinno, K. Kamada, A. A. Starobinsky and J. Yokoyama, JCAP 2101 (2021) 066 [arXiv:2007.10369]

Another mechanism of rapid creation and heating of matter after inflation: tachyonic instability of the Higgs field leading to formation of quasi-classical matter inhomogeneity. It arises when the background Higgs field stays long near $\chi = 0$ in the regime $\phi > 0$.

Occurs not for all values of parameters M, ξ and requires some fine-tuning of them to be efficient: at least $< \mathcal{O}(0.1)$ in the deep Higgs-like regime with a large scalaron mass, while more severe fine-tuning $\sim \mathcal{O}(10^{-4}) - \mathcal{O}(10^{-5})$ is needed in the R^2-like regime with a small non-minimal coupling.

Stochastic behaviour of ϕ, R and χ in this regime – stochastic reheating.



Upper left: $\xi = 4000$. Upper right: $\xi = 4100$. Lower left: $\xi = 4435.759104801013$. Lower right: $\xi = 4435.7591048$. All the digits shown above are needed.

Viable two-parametric inflationary models in GR

The simplest inflationary models with a scalar field minimally coupled to gravity which are not excluded by observations are two-parametric. Their potential can be derived assuming the "no-scale" behaviour $P_{\mathcal{R}}(k) \propto (N+1)^{\beta}$ and using the reconstruction method.

In the slow-roll approximation:

$$\frac{V^3}{V'^2} = C P_{\mathcal{R}}(k(t(\phi))), \quad C = \frac{12\pi^2}{\kappa^6}$$

Changing variables for ϕ to $N(\phi)$ and integrating, we get:

$$\frac{1}{V(N)} = -\frac{\kappa^4}{12\pi^2} \int \frac{dN}{P_{\mathcal{R}}(N)}$$

$$\kappa\phi = \int dN \sqrt{\frac{d \ln V}{dN}}$$

Here, $N \gg 1$ stands both for $\ln(k_f/k)$ at the present time and the number of e-folds back in time from the end of inflation. First derived in H. M. Hodges and G. R. Blumenthal, Phys. Rev. D 42, 3329 (1990).

The two-parameter family of isospectral slow-roll inflationary models, but the second parameter shifts the field ϕ only. The best fit to the Planck data is produced by $\beta = 2$. Then:

$$V = V_0 \frac{N+1}{N+1+N_0} = V_0 \tanh^2 \frac{\kappa\phi}{2\sqrt{N_0}}$$

$$r = \frac{8N_0}{(N+1)(N+1+N_0)}$$

$r \sim 0.003$ for $N_0 \sim 1$. From the upper limit on $r_{0.002}$: $N_0 < 58$ for $N = 56$.

Perspectives of future discoveries

- Primordial gravitational waves from inflation: r. $r \lesssim 8(1 - n_s) \approx 0.3$ (confirmed!) but may be much less. However, under reasonable assumptions one may expect that $r \gtrsim (n_s - 1)^2 \approx 10^{-3}$. The target prediction in the simplest (one-parametric) models is $r = 3(n_s - 1)^2 \approx 0.004$.
- A more precise measurement of $n_s - 1 \Longrightarrow$ duration of transition from inflation to the radiation dominated stage $\Longrightarrow$ information on inflaton (scalaron) couplings to known elementary particles at super-high energies $E \lesssim 10^{13}\ \mathrm{Gev}$.
- Local non-smooth features in the scalar power spectrum at cosmological scales (?).
- Local enhancement of the power spectrum at small scales leading to a significant amount of primordial black holes (?).

Generating peaks and troughs in the primordial scalar spectrum

To obtain large peaks and troughs in $P_{\mathcal{R}}$, temporal breaking of the slow-roll approximation during inflation is needed. The simplest way: fast break in the first derivative of the inflaton potential $V(\phi)$ (A. A. Starobinsky, JETP Lett. 55, 489 (1992)). Leads to a step in $P_{\mathcal{R}}$ with superimposed oscillations. To obtain a peak, two such features with opposite signs, or a fast break in the $V(\phi)$ itself are needed (so that an inflection point appears in between). However, it is not sufficient to have an inflection point only, it should be combined with a strong breaking of the slow-roll conditions.

Let $V(\phi) = V_0 + A_+\phi\,\theta(\phi-\phi_0) + A_-\phi\,\theta(\phi_0-\phi)$ for ϕ close to ϕ_0. Then

$$\dot{\phi} = -\frac{A_+}{3H_0}\,\theta(-t) - \frac{A_- + (A_+ - A_-)e^{-3H_0 t}}{3H_0}\,\theta(t)$$

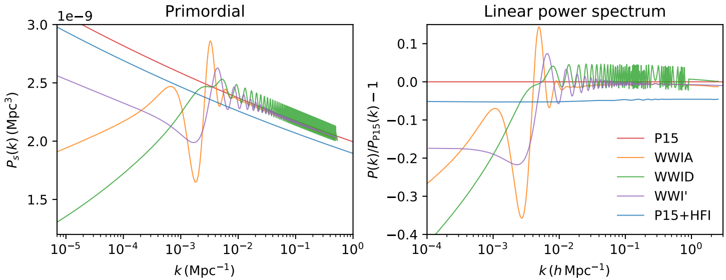
The slow-roll spectrum $P_{\mathcal{R}}$ is modulated by the multiplier

$$D^2 = 1 - 3\left(\frac{A_-}{A_+} - 1\right)\left[\left(1 - \frac{1}{y^2}\right)\sin 2y + \frac{2}{y}\cos 2y\right] +$$

$$\frac{9}{2}\left(\frac{A_-}{A_+} - 1\right)^2 \frac{1}{y^2}\left(1 + \frac{1}{y^2}\right) \times$$

$$\left[1 + \frac{1}{y^2} + \left(1 - \frac{1}{y^2}\right)\cos 2y - \frac{2}{y}\sin 2y\right],$$

$$y = \frac{k}{k_0},\ D(0) = \frac{A_-}{A_+},\ D(\infty) = 1$$


Primordial
1e−9
3.0
2.5
2.0
1.5
$P_s(k)\,(\mathrm{Mpc}^3)$
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
$k\,(\mathrm{Mpc}^{-1})$
Linear power spectrum
0.1
0.0
−0.1
−0.2
−0.3
−0.4
$P(k)/P_{\mathrm{P15}}(k) - 1$
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
$k\,(h\,\mathrm{Mpc}^{-1})$
P15
WWIA
WWID
WWI'
P15+HFI

PBHs and small-scale GWs in two-field models of inflation

M. Braglia, D. K. Hazra, F. Fimelli, G. F. Smoot,
L. Sriramkumar, A. A. Starobinsky, JCAP 2008 (2020) 001.

The simplest one-parameter inflationary models do not predict PBHs, at least whose existing at present with $M > 10^{15}$ g.
Previously known ways to obtain a large peak in the primordial power spectrum of scalar adiabatic perturbations at small scales:
1. A local feature in the inflaton potential $V(\phi)$ (a rapid change of its slope or its amplitude, an inflection point with a large $V'''(\phi)$).
2. A rapid turn of the inflaton trajectory in the field space in the case of many-field models of inflation.
3. Phase transitions leading to large isocurvature perturbations which transform to adiabatic ones afterwards.

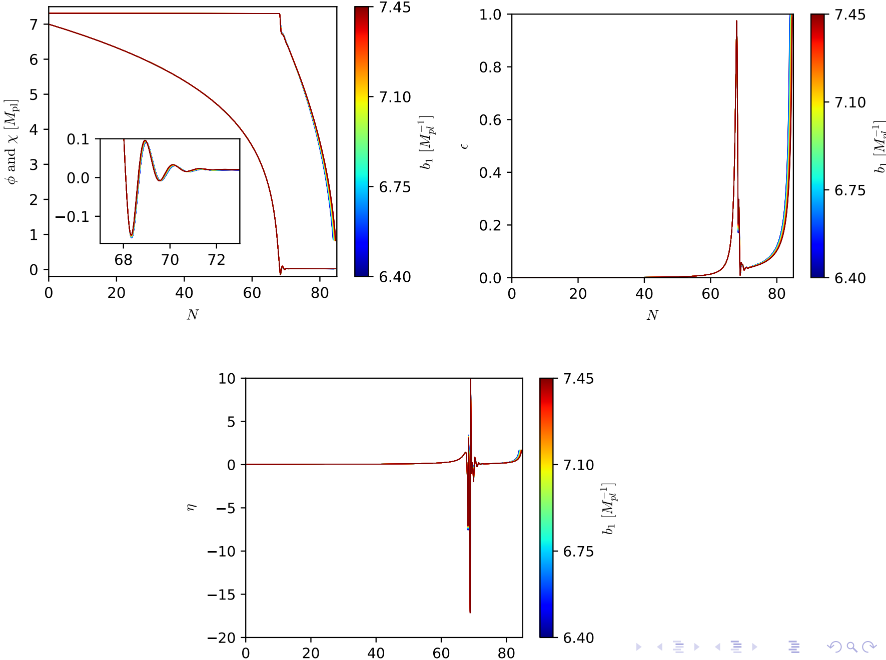
Two-field inflation with large kinetic coupling

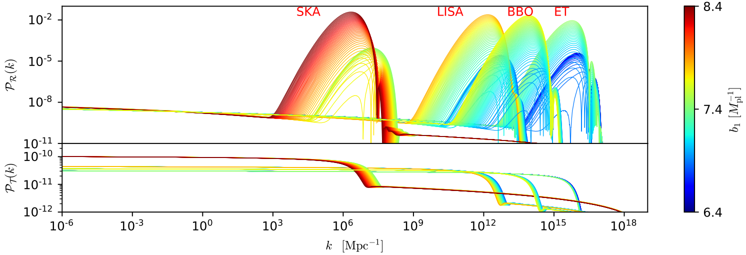
A novel mechanism: a two-field inflation with different inflaton effective masses (that leads to two stages of inflation) and a large non-standard kinetic coupling of the heavier field.

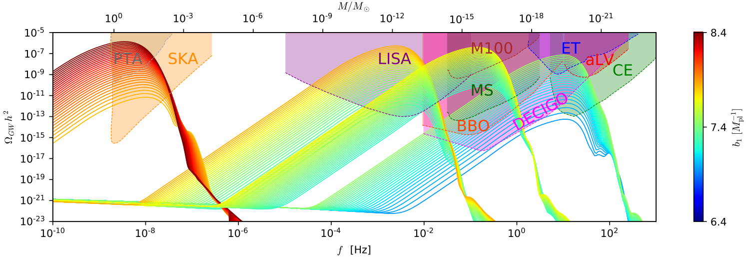
$$S(\phi,\chi) = \int d^4x\,\sqrt{-g}\,\left[\frac{M_{Pl}^2}{2}R - \frac{1}{2}(\partial\phi)^2 - \frac{f(\phi)}{2}(\partial\chi)^2 - V(\phi,\chi)\right]$$

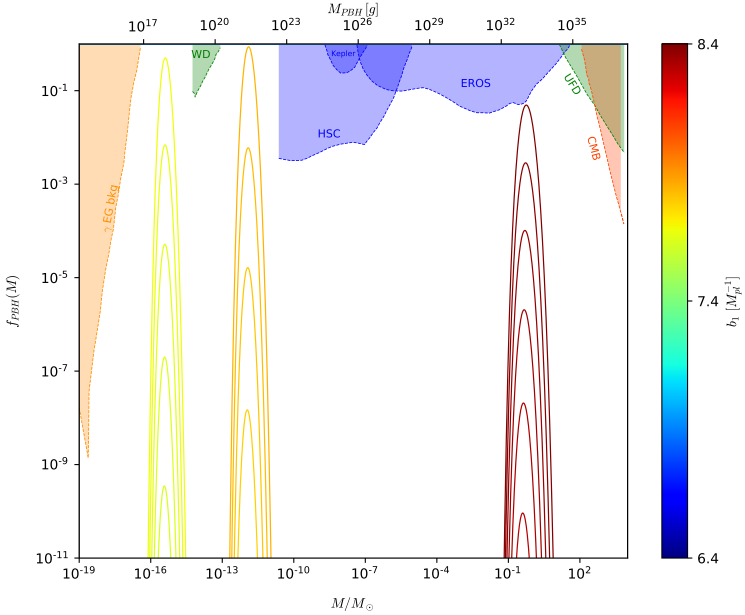
$$V(\phi,\,\chi) = V_0\frac{\phi^2}{\phi_0^2+\phi^2} + \frac{m_\chi^2}{2}\chi^2$$

Large kinetic coupling: $f(\phi) = \exp(b\phi),\ bM_{Pl} \gg 1$. The peak in the spectrum arises when the heavier field goes out of the slow-roll regime. It can lead to the formation of PBHs with a wide range of masses and to the generation of stochastic background of primordial gravitational waves produced by second order scalar perturbations.


ϕ and χ $[M_{\rm pl}]$
N
b_1 $[M_{pl}^{-1}]$
7.45
7.10
6.75
6.40
ϵ
N
η


SKA
LISA
BBO
ET
$\mathcal{P}_{\mathcal{R}}(k)$
$\mathcal{P}_{\mathcal{T}}(k)$
k [Mpc^{-1}]
b_1 [$M_{\rm pl}^{-1}$]
8.4
7.4
6.4


$M/M_\odot$
PTA
SKA
LISA
M100
ET
aLV
CE
MS
BBO
DECIGO
$\Omega_{GW}h^2$
f [Hz]
b_1 [$M_{\rm pl}^{-1}$]
8.4
7.4
6.4



$M_{PBH}\,[g]$
10^{17}
10^{20}
10^{23}
10^{26}
10^{29}
10^{32}
10^{35}
WD
Kepler
EROS
HSC
UFD
CMB
γ EG bkg
$f_{PBH}(M)$
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
10^{-19}
10^{-16}
10^{-13}
10^{-10}
10^{-7}
10^{-4}
10^{-1}
10^{2}
$M/M_\odot$
8.4
7.4
6.4
$b_1\,[M_{pl}^{-1}]$

Generality of inflation

Theorem. In inflationary models in GR and $f(R)$ gravity, there exists an open set of classical solutions with a non-zero measure in the space of initial conditions at curvatures much exceeding those during inflation which have a metastable inflationary stage with a given number of e-folds.

For the GR inflationary model this follows from the generic late-time asymptotic solution for GR with a cosmological constant found in A. A. Starobinsky, JETP Lett. 37, 55 (1983). For the $R + R^2$ model, this was proved in A. A. Starobinsky and H.-J. Schmidt, Class. Quant. Grav. 4, 695 (1987). For the power-law and $f(R) = R^p, p < 2$, $2 - p \ll 1$ inflation – in V. Müller, H.-J. Schmidt and A. A. Starobinsky, Class. Quant. Grav. 7, 1163 (1990).

Generic late-time asymptote of classical solutions of GR with a cosmological constant Λ both without and with hydrodynamic matter (also called the Fefferman-Graham expansion):

$$ds^2 = dt^2 - \gamma_{ik} dx^i dx^k$$

$$\gamma_{ik} = e^{2H_0 t} a_{ik} + b_{ik} + e^{-H_0 t} c_{ik} + \ldots$$

where $H_0^2 = \Lambda/3$ and the matrices a_{ik}, b_{ik}, c_{ik} are functions of spatial coordinates. a_{ik} contains two independent physical functions (after 3 spatial rotations and 1 shift in time + spatial dilatation) and can be made unimodular, in particular. b_{ik} is unambiguously defined through the 3-D Ricci tensor constructed from a_{ik}. c_{ik} contains a number of arbitrary physical functions (two - in the vacuum case, or with radiation) – tensor hair.

A similar but more complicated construction with an additional dependence of H_0 on spatial coordinates in the case of $f(R) = R^p$ inflation – scalar hair.

Consequences:

1. (Quasi-) de Sitter hair exist globally and are partially observable after the end of inflation.

2. The appearance of an inflating patch does not require that all parts of this patch should be causally connected at the beginning of inflation.

Similar property in the case of a generic curvature singularity formed at a spacelike hypersurface in GR and modified gravity. However, 'generic' does not mean 'omnipresent'.

Pre-inflationary stage

Different possibilities were considered historically:
1. Creation of inflation "from nothing" (Grishchuk and Zeldovich 1981).
One possibility among infinite number of others.
2. De Sitter "Genesis": beginning from the exact contracting full de Sitter space-time at $t \to -\infty$ (AS 1980).
Requires adding an additional term

$$R_i^l R_l^k - \frac{2}{3} R R_i^k - \frac{1}{2} \delta_i^k R_{lm} R^{lm} + \frac{1}{4} \delta_i^k R^2$$

to the rhs of the gravitational field equations. Not generic.
May not be the "ultimate" solution: a quantum system may not spend an infinite time in an unstable state.
3. Bounce due to a positive spatial curvature (AS 1978).
Generic, but probability of a bounce is small for a large initial size of a universe $W \sim 1/Ma_0$. It is difficult to reach inflation from a low curvature state.

Formation of inflation from generic curvature singularity

In classical gravity (GR or modified $f(R)$): space-like curvature singularity is generic. Generic initial conditions near a curvature singularity in modified gravity models (the $R + R^2$ and Higgs ones): anisotropic and inhomogeneous (though quasi-homogeneous locally). Recent analytical and numerical investigation for $f(R)$ gravity in the Bianchi I type model in D. Muller, A. Ricciardone, A. A. Starobinsky and A. V. Toporensky, Eur. Phys. J. C **78**, 311 (2018). Two types of singularities in with the same structure at $t \to 0$:

$$ds^2 = dt^2 - \sum_{i=1}^{3} |t|^{2p_i}\, a_l^{(i)} a_m^{(i)}\, dx^l dx^m,\ 0 < s \leq 3/2,\ u = s(2-s)$$

where $p_i < 1,\ s = \sum_i p_i,\ u = \sum_i p_i^2$ and $a_l^{(i)},\ p_i$ are functions of $\mathbf{r}$. It is easy to reach inflation from a state with curvature exceeding that during the last 55 e-folds of inflation.

Bianchi I type models with inflation in R^2 gravity

Type A. $1 \leq s \leq 3/2,\ R \propto |t|^{1-s} \to +\infty$
Type B. $0 < s < 1,\ R \to R_0 < 0,\ f'(R_0) = 0$

For $f(R) = R^2$ even an exact solution can be found.

$$ds^2 = \tanh^{2\alpha}\left(\frac{3H_0 t}{2}\right)\left(dt^2 - \sum_{i=1}^{3} a_i^2(t)dx_i^2\right)$$

$$a_i(t) = \sinh^{1/3}(3H_0 t)\tanh^{\beta_i}\left(\frac{3H_0 t}{2}\right),\ \sum_i \beta_i = 0,\ \sum_i \beta_i^2 < \frac{2}{3}$$

$$\alpha^2 = \frac{\frac{2}{3} - \sum_i \beta_i^2}{6},\quad \alpha > 0$$

Next step: relate arbitrary functions of spatial coordinates in the generic solution near a curvature singularity to those in the quasi-de Sitter solution. Spatial gradients may become important for some period before the beginning of inflation.

The same structure of generic singulariry for a non-minimally coupled scalar field (scalar-tensor gravity) (as well as for the mixed Higgs-R^2 model):

$$S = \int \left(f(\phi)R + \frac{1}{2}\phi_{,\mu}\phi^{,\mu} - V(\phi) \right) \sqrt{-g}\, d^4x + S_m$$

$$f(\phi) = \frac{1}{2\kappa^2} - \xi\phi^2$$

Type A. $\xi < 0, |\phi| \to \infty$
Type B. $\xi > 0, |\phi| \to 1/\sqrt{2\xi}\kappa$

The asymptotic regimes and a number of exact solutions in the Bianchi type I model are presented in A. Yu. Kamenshchik, E. O. Pozdeeva, A. A. Starobinsky, A. Tronconi, G. Venturi and S. Yu. Vernov, Phys. Rev. D 97, 023536 (2018) with some of them borrowed from A. A. Starobinsky, MS Degree thesis, Moscow State University, 1971, unpublished.

What is sufficient for beginning of inflation in classical (modified) gravity, is:
1) the existence of a sufficiently large compact expanding region of space with the Riemann curvature much exceeding that during the end of inflation ($\sim M^2$) – realized near a curvature singularity;
2) the average value $< R >$ over this region positive and much exceeding $\sim M^2$, too, – type A singularity;
3) the average spatial curvature over the region is either negative, or not too positive.

On the other hand, causal connection is certainly needed to have a "graceful exit" from inflation, i.e. to have practically the same amount of the total number of e-folds during inflation N_{tot} in some sub-domain of this inflating patch.

Conclusions

- The typical inflationary predictions that $|n_s - 1|$ is small and of the order of N_H^{-1}, r does not exceed $\sim 8(1 - n_s)$ and initial perturbations are Gaussian with high accuracy has been confirmed. Typical consequences following without assuming additional small parameters: $H_{55} \sim 10^{14}\,\mathrm{GeV}$, $m_{infl} \sim 10^{13}\,\mathrm{GeV}$.
- In $f(R)$ gravity, the simplest $R + R^2$ model is one-parametric and has the preferred values $n_s - 1 = -\frac{2}{N} \approx -0.035$ and $r = 3(n_s - 1)^2 \approx 0.004$. The first value produces the best fit to present observational CMB data. The same prediction follows for the Higgs and the mixed R^2-Higgs models though actual values of N are slightly different for these 3 cases.
- Inflation in $f(R)$ gravity represents a dynamical attractor for slow-rolling scalar fields strongly coupled to gravity.

- The mixed R^2-Higgs model helps to remove some UV problems of the Higgs inflationary model and may be considered as its UV-completion up to the Planck energy.
- The rate of post-inflationary heating though particle creation in the mixed R^2-Higgs model is intermediate between those in the Higgs and $R + R^2$ models. Generically inflaton (scalaron) decay is not instant and occurs after a large number of its oscillations.
- In some fine-tuned sub-regions of parameters of the mixed R^2-Higgs model, more rapid preheating through tachyonic instability of the Higgs field becomes possible.

- In two-field inflationary models with a large non-standard kinetic coupling of a heavier inflaton field, it is possible to produce a large peak at small scales in the primordial power spectrum of scalar adiabatic perturbations leading to the formation of PBHs and to the peak in the stochastic background of primordial gravitational waves produced by second order scalar perturbations.
- Inflation is generic in the $R + R^2$ inflationary model and close ones. Thus, its beginning does not require causal connection of all parts of an inflating patch of space-time (similar to space-like singularities). However, graceful exit from inflation requires approximately the same number of e-folds during it for a sufficiently large compact set of geodesics. To achieve this, causal connection inside this set is necessary (though still may appear insufficient).
- Inflation can form generically and with not a small probability from generic space-like curvature singularity.