N-formalism and beyond

Alexei A. Starobinsky

Landau Institute for Theoretical Physics RAS,
Moscow and
RESCEU, The University of Tokyo, Tokyo

Nishinomiya Yukawa Symposium: New Waves in Gravity and Cosmology
Celebrating the 60th birthday of Prof. Misao Sasaki
Kyoto, Japan, 04.12.2012

δN formalism

Observing ΔN

Going to the non-linear case – N-formalism

N-formalism in the regime of stochastic inflation

Conclusions

δN formalism

Aim: relate large-scale inhomogeneities in the remote past – soon after the first Hubble radius $r_H = H^{-1}(t) = a(t)/\dot{a}(t)$ crossing in the early Universe – to those at the present time.

General idea: as far as as $\lambda \gg r_H$, there is no both-way causal contact between points separated by spatial distances $\sim \lambda$ and more, so:
a) spatial gradients can be neglected;
b) evolution of perturbations does not depend on local physics (in fact, some of them remain constant).

Works not for all perturbations, but for at least 3 modes of them: the constant scalar ('growing' adiabatic) mode and two polarizations of the constant tensor (GW) mode.

For the most interesting case of the adiabatic scalar mode:

$$\zeta(\mathbf{r}) = \Delta N(t_{in}, t_{fin}(\mathbf{r})) \equiv \Delta \ln \frac{a(t_{fin})}{a(t_{in})} = \sum_a \frac{\delta N}{\delta \phi_a} \delta\phi_a(\mathbf{r}, t_{in})$$

where $\frac{\delta N}{\delta \phi_a}$ is estimated using a homogeneous spatially flat FRW solution for any matter content, t_{in} is taken soon after the first Hubble radius crossing and t_{fin} is taken at the moment when the complete reheating occurs or even at the recombination time.

First appeared (for the one-inflaton, but fully non-linear case) in A. A. Starobinsky, Phys. Lett. B 117, 175 (1982), then it was generalized for a specific (linear) case of multiple inflation (when $V = \sum_a V_a(\phi_a)$) in A. A. Starobinsky, Sov. Astron. Lett. 11, 271 (1985), and then to an arbitrary multiple inflation in M. Sasaki and E. D. Stewart, Prog. Theor. Phys. 95, 71 (1996). After that Sasaki-san wrote many papers on this topic, the last one is A. Naruko, Y. Takamizu and M. Sasaki, arXiv:1210.6525.

What is needed and what is not needed

What is needed for this expression to be true and have sense:
1) the existence of spatially-flat FRW solutions, spatial curvature can be neglected through the whole evolution between t_{in} and t_{fin};
2) the existence of the first Hubble radius crossing;
3) slow-roll conditions at t_{in} (in particular, in order to calculate $\delta\phi_a$);
4) decaying modes can be neglected through the whole evolution (in particular, this may not be possible around the moment when $H = 0, \dot{H} \neq 0$);
5) not too steep power spectrum of scalar perturbations (certainly, $n_s < 5$).

What is not needed:
1) slow-roll after t_{in};
2) straightness of an evolution trajectory in the field space in the case of multiple scalar fields.

What is covered

Which cases are covered by this expression:
1) usual fluctuations in multiple inflation;
2) curvaton fluctuations after inflation: e.g. for a massive curvaton σ with $mt_{reh} \gg 1$ and $t_d \gg t_{eq}$,

$$N = N_0 + \frac{1}{2} \ln \frac{t_{eq}}{t_{reh}} + \frac{2}{3} \ln \frac{t_d}{t_{eq}}, \quad t_{eq} \propto \sigma^{-4}$$

so

$$\zeta = \left(\frac{1}{2} - \frac{2}{3} \right) \frac{\delta t_{eq}}{t_{eq}} = \frac{2}{3} \frac{\delta \sigma}{\sigma}$$

3) modulated decay of an inflaton:

$$N = N_0 + \frac{2}{3} \ln \frac{H_f}{\Gamma} + \frac{1}{2} \ln (\Gamma t_{reh})$$

$$\zeta = -\frac{1}{6} \frac{\delta \Gamma}{\Gamma} = -\frac{1}{6} \frac{\delta \ln \Gamma}{\delta \chi} \delta \chi$$

4) fluctuations from the end of inflation, etc.

Observing ΔN

The simplest way: by looking at the large-angle ($l \lesssim 50$) CMB fluctuations.
Neglecting small contribution from radiation to the expansion rate at the moment of recombination and the small ISW effect,

$$\frac{\delta T(\theta, \phi)}{T} = \frac{1}{3}\Phi(r_{lss}, \theta, \phi) = -\frac{1}{5}\left(N(r_{lss}, \theta, \phi) - < N >_{lss}\right)$$

where the averaging is made over the last scattering surface.
In particular, the well known cold spot in the $\delta T/T$ map can be interpreted as a local maximum in $N(\mathbf{r})$.

Going to the non-linear case– N-formalism

In the synchronous system of reference with the additional conditions that t_{in} does not depend on $\mathbf{r}$ and that the spatial metric is conformally flat at large scales initially,

$$ds^2 = dt^2 - a^2(t)e^{2\zeta(\mathbf{r})}d\mathbf{r}^2 + \text{small terms}$$

where

$$\zeta(\mathbf{r}) = N(\mathbf{r})$$

and N is calculated as a functional of initial (homogeneous) values of $\phi_{a,in}$ using a spatially flat FRW solution.
Thus, N can be considered as a generating functional for the growing mode of inhomogeneities (not for all possible inhomogeneities!). Then it can be expanded in powers of $\delta\phi_{a,in}$ – homogeneous deviations from a chosen background solution – for calculation of non-linear corrections including non-Gaussianity.

Very simple derivation

Following my PLB 1982 paper. Let $H \approx H_0$ during inflation for simplicity, and let inflation begins at $t = 0$ and ends at $t = t_f(\mathbf{r})$. Then the initial unperturbed FRW metric is

$$ds^2 = dt^2 - a_0^2 e^{2H_0 t} d\mathbf{r}^2 \equiv dt^2 - a_0^2 e^{2H_0 t_f(\mathbf{r})} e^{2H_0(t - t_f(\mathbf{r}))} d\mathbf{r}^2$$

As a result of quantum fluctuations,

$$\phi(t) \to \phi(t - t_f(\mathbf{r})), \; a(t) \to a(t - t_f(\mathbf{r}))$$

So, the perturbed metric further evolves as

$$ds^2 = dt^2 - a_0^2 e^{2H_0 t_f(\mathbf{r})} a^2(t - t_f(\mathbf{r})) d\mathbf{r}^2 + \text{small terms}$$

that finally, for $t \gg t_f$, takes the form

$$ds^2 = dt^2 - a_0^2 e^{2H_0 t_f(\mathbf{r})} a^2(t) d\mathbf{r}^2 + \text{small terms}$$

Quasi-isotropic solution

This derivation uses a particular form of a more general statement:

If a metric theory of gravity with any matter (which may be multicomponent) admits a spatially flat FRW solution with the scale factor $a(t)$, then there exists its inhomogeneous generalization with at least 3 arbitrary physical functions of spatial coordinates which in the limit $\lambda \gg r_H$ has the form

$$ds^2 = dt^2 - a^2(t)A_{ik}dx^i dx^k + \mathcal{O}\left(\left(\frac{r_H}{\lambda}\right)^2\right), \; i, k = 1, 2, 3$$

where the matrix A_{ik} arbitrarily (but with bounded derivatives) depends on spatial coordinates x^i.

The matrix A_{ik} contains 3 physical functions remaining after 3 coordinate transformations – just the required number.

This solution was called quasi-isotropic by I. M. Khalatnikov and E. M. Lifshitz who first considered it in the particular, one-component case of radiation as early as in 1960.

Constant adiabatic mode in generic multiple inflation

Another consequence of this general statement:

For any multicomponent inflation, there always exists one solution for scalar perturbations for which $\zeta = const$ in the limit $k \to 0$.

See, e.g., such explicit statement in A. A. Starobinsky, S. Tsujikawa and J. Yokoyama, Nucl. Phys. B 610, 383 (2001) [arXiv:astro-ph/0107555]. It was also implicitly used in my Sov. Astron. Lett. 1985 paper.
It is natural to call this solution a constant ('growing') adiabatic mode. This should not be confused with the popular way to define an 'adiabatic perturbation' by projecting scalar perturbations in multiple inflation on the unperturbed trajectory in the field space. However, the latter way hides the existence of this simple constant solution.

Effective quantum-to-classical transition and account of large fluctuations

Since field fluctuations $\delta\phi_a$ are quantum, the resulting space-time metric is quantum, too, even at present. This does not lead to any problems for the N-formalism since if a decaying mode can be neglected, the resulting field operators commute with everything. This makes possible the introduction of their effective c-number, though stochastic description.

In the case of slow-roll inflation, the effect of large though sufficiently smooth stochastic field fluctuations on local expansion can be taken into account non-perturbatively using the stochastic inflation approach.

Necessary steps (single inflation is considered below for simplicity):

I. Solving the Fokker-Planck equation for the one-point probability distribution $\rho(\phi, \tau)$ of an inflaton scalar field ϕ during slow-roll inflation in terms of the time $\tau = \int Hdt = \ln a = N$.

$$\frac{\partial \rho}{\partial \tau} = \frac{\partial}{\partial \phi}\left(\frac{V'}{3H^2}\,\rho\right) + \frac{1}{8\pi^2}\,\frac{\partial^2}{\partial \phi^2}\left(H^2\rho\right)$$

Probability conservation: $\int \rho\, d\phi = 1$. Of course, the solution will depend on an initial condition $\rho_0(\tau)$ at the local beginning of inflation.

II. Calculating n-point probability distributions.
Following A.A. Starobinsky and J. Yokoyama, Phys. Rev. D **50**, 6357 (1994).
In the leading approximations, all Green functions and joint n-point probability distributions of the inflaton field can be expressed through solutions of the same Fokker-Planck equation with different initial conditions only. In particular, in the case $H \approx H_0$ during inflation (for simplicity), the general two-point PDF for 4D-points lying outside each other future light cones in the stochastic approach is:

$$\rho_2[\phi_1(\mathbf{r_1}, t_1), \phi_2(\mathbf{r_2}, t_2)] =$$

$$\int \Pi[\phi_1(\mathbf{r_1}, t_1)|\phi_r(\mathbf{r_1}, t_r)]\Pi[\phi_2(\mathbf{r_2}, t_2)|\phi_r(\mathbf{r_2}, t_r)]\rho_1(\phi_r, t_r)\, d\phi_r$$

where t_r is the time in the past when both corresponding 3D-points were inside one Hubble volume and $\Pi[\phi_1(\mathbf{r}, t_1)|\phi_2(\mathbf{r}, t_2)]$ satisfies the Fokker-Planck equation with respect to both its time and field variables with the initial condition

$$\Pi[\phi_1(\mathbf{r}, t_1)|\phi_2(\mathbf{r}, t_1)] = \delta(\phi_1 - \phi_2)$$

Otherwise, if the 4D-points are inside the future light cone of one of them, the spatial points $\mathbf{r_1}$ and $\mathbf{r_2}$ are inside one elementary averaging volume, so they coincide in terms of the stochastic approach. Then, for $t_1 < t_2$,

$$\rho_2[\phi_1(\mathbf{r}, t_1), \phi_2(\mathbf{r}, t_2)] = \Pi[\phi_2(\mathbf{r}, t_2)|\phi_1(\mathbf{r}, t_1)]\rho_1(\phi_1, t_1))$$

III. Making the transition to predictions for the post-inflationary evolution.

From $\rho(\phi, \tau)$ during inflation to the distribution $w(\tau)$ over the total local duration of inflation:

$$w(\tau) = \lim_{\phi \to \phi_{end}} j = \lim_{\phi \to \phi_{end}} \frac{|V'|}{3H^2} \rho(\phi, \tau)$$

For the graceful exit to a post-inflationary epoch, the stochastic force should be much less than the classical one during last e-folds of inflation.

The same way to obtain the joint distribution $w(0, \tau_1;\ |\mathbf{r}|, \tau_2)$ from the 2-point joint probability distribution $\rho(\phi_1, 0, \tau_1; \phi_2, |\mathbf{r}|, \tau_2)$ during inflation.

Conclusions

- δN-formalism and its non-linear generalization N-formalism work simply and successfully for a wide class of metric theories of gravity and different types of matter.
- However, since they refer to some special class of solutions only, their applicability for a given particular background solution should be always checked.
- Using the stochastic inflation approach, the N-formalism can be used in the regime of large though smooth deviations from isotropy (large ΔN).

- Three ways to go beyond the N-formalism:
  a) perturbative account of small terms in powers of $\sim (r_H/\lambda)^2$;
  b) account of primordial GW and their non-linear effects on scalar perturbations;
  c) perturbative account of small decaying modes.
  However, in the latter case the quantum nature of perturbations may show itself.

So, the whole topic is flourishing and experiences the second birth due to its successful application in completely new situations, and I wish the same to Sasaki-san!