Getting the Measure of the Flatness Problem

By **Guillaume Evrard**[1,2] **and Peter Coles**[2]

[1] Groupe de Recherche en Astronomie et Astrophysique du Languedoc,

URA 1368 CNRS/Montpellier II, c.c. 072, Université Montpellier II,

place Eugène Bataillon, F-34095 Montpellier Cedex 05, France

[2] Astronomy Unit, School of Mathematical Sciences,

Queen Mary and Westfield College, Mile End Road,

London E1 4NS, United Kingdom

Summary

The problem of estimating cosmological parameters such as Ω from noisy or incomplete data is an example of an inverse problem and, as such, generally requires a probablistic approach. We adopt the Bayesian interpretation of probability for such problems and stress the connection between probability and information which this approach makes explicit. This connection is important even when information is "minimal" or, in other words, when we need to argue from a state of maximum ignorance. We use the transformation group method of Jaynes to assign minimally–informative prior probability measure for cosmological parameters in the

1

simple example of a dust Friedman model, showing that the usual statements of the cosmological flatness problem are based on an inappropriate choice of prior. We further demonstrate that, in the framework of a classical cosmological model, there is no flatness problem.

In the physical sciences, the word "model" is usually used to denote a theoretical description of a system that contains one or more "free parameters" whose values can not be determined *a priori* but which have to be estimated by empirical means. Such estimation problems generally go under the name of "inverse problems" and, because available data are often incomplete or noisy, they generally require probabilistic reasoning.

Modern 'Big Bang' cosmology rests on a mathematical framework supplied by the simplest relativistic cosmological models compatible with the Cosmological Principle, i.e. the Friedman models. These models have two free parameters, the Hubble parameter, H_0, and the deceleration parameter q_0 (or, equivalently for these models, the deceleration parameter $q_0 = \Omega_0/2$; the suffix "0" indicates that the parameter in question is measured at the present epoch, i.e. when the cosmological proper time is t_0.) As is the case for physical models in general, these parameters are not *predicted* by the Big Bang theory itself, but need to be inferred from observational data. Because the values of H and Ω at any time can be determined from the present values H_0 and Ω_0 if the model is specified, it is in principle possible to learn about conditions very near the Big Bang singularity from estimates of the cosmological parameters made at the present time.

2

The problem with Ω is that its value is not known with any precision: it probably lies in the range $0.10 < \Omega_0 < 1.5$, but the relevant evidence is often contradictory[1]. However, Ω evolves strongly with cosmic time t in such a way that $\Omega = 1$ is an unstable fixed point. To get a value of Ω anywhere near unity at the present time (even a factor of a few either way) consequently requires a value at very early times extremely close to unity (say $\Omega = 1 \pm 10^{-60}$ at the Planck time). The *cosmological flatness problem* arises from the judgement that this "fine–tuning" is somehow unlikely on the basis of standard Friedman models; it is is usually "resolved" by appealing to some transient mechanism (e.g. inflation[2]) which can make Ω evolve towards unity for some time, rather than away from it.

But do we have any right to claim that some values of Ω are more likely than others? Can one make any inferences at all from the uncertain parameter estimates we have in cosmology? And what precisely does it mean to say that Ω is "close to unity" anyway?

To answer these questions we need to understand the role of probability in the solution of inverse problems generally[3]. We adopt the objective Bayesian interpretation of probability which, we believe, is the only way to formulate this type of reasoning in a fully self– consistent way. In this interpretation, probability represents a generalisation of the notions of "true" and "false" to intermediate cases where there is insufficient information to decide with logical certainty between these two alternatives[4]. Unlike the opposing "frequentist" view, the Bayesian lends itself naturally to the interpretation of unique events, of which the Big Bang is the most obvious relevant example[5].

3

The central principle involved in Bayesian inference is Bayes' theorem[6]. Suppose H_i represents one of a set of hypotheses (or models), D is some data and I is whatever relevant prior information we may have (or which we assume to be the case) before obtaining the data D. Bayes' theorem states that

$$P(H_i|DI) = \frac{P(H_i|I)P(D|H_iI)}{\sum_i P(H_i|I)P(D|H_iI)}, \tag{1}$$

where $P(H_i|I)$ is called the prior probability of H_i given our prior information, $P(D|H_iI)$ is the likelihood and $P(H_i|DI)$ is the posterior probability. Notice that all probabilities here are conditional on the information I which is either known or assumed to be true in a given model. If the prior is relatively flat and the likelihood of the data D is strongly peaked for a particular H_i then our inference of the posterior probability is strongly determined by the data. If, on the other hand, the data discriminate only weakly between the models then the posterior is dominated by the prior. In general, however, both prior and likelihood are required for the inverse problem to be well–posed. Many critics have dubbed the Bayesian approach "subjective" because different individuals may possess different information and therefore assign different priors to the same hypothesis. This is not a serious objection: your assessment of the probability that a given horse will win a race must change if you learn the other horses have all been drugged! What is important is that, given the same information, the same prior should be assigned. We therefore need an objective set of rules for assigning priors when information is specified. In particular, we may have no information at all other than that inherent in the model we adopt. What should one do when one has such minimal information about a system?

4

Even this apparently simple question turns out to be extremely deep and there is no universally accepted principle for assigning minimally–informative priors in general circumstances. Jaynes[7] has described one approach which is, as far as we are aware, the most general objective algorithm available. "Jaynes' principle" is that one looks for a measure on the parameter space of the system that possesses the property of invariance under the group of transformations which leave unchanged the mathematical form of the physical laws describing the system. In the absence of any other constraints, the principle of maximum information entropy (a principle of least prejudice) yields a prior probability simply proportional to this measure.

To take a trivial illustrative example, consider the problem of estimating the position of a particle on the real line. Our state of knowledge, if no signposts are visible, must be unchanged if we shift our coordinates by any distance γ. This requires $\mu(x) = \mu(x + \gamma)$, a functional equation which has only one solution: $\mu =$ constant. This is in full accord with our intuition, but it does not mean that a uniform prior is appropriate for all cases where we are seeking to encode minimal information. For example, Evrard[8] has calculated the least–informative prior for a free particle in velocity space using Jaynes' principle and the laws of special relativity. Even in this simple example, the result is non–trivial: "least information prior" does not necessarily mean "no prior".

We now turn to the appropriate minimally informative prior for the cosmological parameters H_0 and Ω_0. We take the laws of physics to be the Friedman equations describing a pressureless perfect fluid in the form

$$a\left(k + \frac{\dot{a}^2}{c^2}\right) = \chi, \qquad (2)$$

5

where χ remains constant throughout the evolution of the system; its value is determined by the "initial value equation"

$$\chi = \frac{4\pi G \rho a^3}{3c^2}.$$ (3)

The quantity χ can be thought of as an absolute scale parameter. In equations (2) & (3), a is the cosmic scale factor (another scale parameter) and ρ is the matter density. The quantity k appearing in equation (2) is the curvature of spatial sections in the model, scaled to take the values 0 if $\Omega = 1$, -1 if $\Omega < 1$ or $+1$ if $\Omega_0 > 1$. The system can be parametrised completely in terms of χ and a. (In fact, we could equally well have chosen to work with redshift z, cosmological proper time t, conformal time τ, temperature T, or anything else monotonically related to a: the resulting measure would turn out to be the same, but the equations turn out to be simpler in terms of a itself.) We now need to express the cosmological parameters $H = \dot{a}/a$ and $\Omega = 2q = -2a\ddot{a}/\dot{a}^2$ in terms of a and χ. We obtain, for $k = \pm 1$,

$$\Omega = 2(2 \mp a/\chi)^{-1}$$ (4)

and

$$H = \left(\frac{c}{\chi}\right) \frac{\sqrt{2 \mp a/\chi}}{(a/\chi)^{3/2}}.$$ (5)

Remember that the suffix 0 represents a quantity defined at the present epoch, so H_0 and Ω_0 are the values of these parameters when $a = a_0$; $\chi = \chi_0$ at all epochs. Because both χ and a are scale parameters, we look for a measure which is invariant under the transformations $a' = \alpha a$ and $\chi' = \beta \chi$, where α and β are constants. Such

6

invariances require that the information represented by our measure does not change if we use a different ruler to measure distances. It follows that

$$\mu(\chi, a) \propto \frac{1}{\chi a},$$

(6)

which becomes, after substituting from equations (4) & (5),

$$\mu(H, \Omega) \propto \frac{1}{H\Omega|\Omega - 1|}.$$

(7).

Note that this measure leads to an *improper* (i.e. non–normalisable) prior probability. This can be rectified by bringing in additional information, such as the ages of cosmic objects which rule out high values of both Ω and H. Anthropic selection effects can also be brought to bear on this question[5]. The measure for H is uniform in the logarithm, as one might expect from the Bayesian "rule of thumb" for scale parameters[9]. The measure in Ω is, however, more complicated than this. In particular, it diverges at $\Omega = 0$ and $\Omega = 1$, the former corresponding to an empty Universe without deceleration and the latter to the critical-density Einstein–De Sitter model. These singularities could have been anticipated because these are two fixed points in the evolution of Ω. A model with $\Omega = 1$ exactly remains in that state forever. Models with $\Omega < 1$ evolve to a state of free expansion with $\Omega = q = 0$. Since states with $0 < \Omega < 1$ are transitory, it is reasonable, in the absence of any other information, to infer that the system should be in one of the two fixed states. (All values of $\Omega > 1$ are transitory.)

The measure (7) also demonstrates how dangerous it is to talk about Ω_0 "near" unity. In terms of our least–informative measure, values of Ω not exactly equal to 1

7

are actually infinitely far from this value. A similar property is held by the velocity–space measure[8], which demonstrates the velocities of all material particles are, in a well–defined sense, infinitely far from c.

We now turn to the flatness problem. The usual argument is essentially that, without inflation, the models that produce $\Omega_0 = 1 \pm \epsilon$ at the present epoch emerge from earlier states with Ω even closer to unity. If one were to adopt a measure which is roughly flat in the vicinity of $\Omega = 1$ as $t \to 0$ then the probability associated with this set of states would vanish and there would indeed be a flatness problem: it would appear "unlikely" that our Universe was correctly modelled by the standard Friedman equations and one would be pushed into accepting inflation as a solution of this "fine–tuning". But our measure (7) demonstrates that the assumption of a constant prior for Ω is not consistent with the assumption of minimal information. It therefore represents a considerable prejudice compared to the least—informative and, therefore, least–prejudiced measure. This prejudice may be motivated to some extent by quantum–gravitational considerations that render the classical model inappropriate, but unless the model adopted and its associated information are stated explicitly one has no right to assign a prior and therefore no right to make any inferences.

Notwithstanding the recent research interest in quantum gravity, we feel that 'minimal knowledge' is a fair description of our state of understanding of physics at the Planck epoch. In terms of the least–informative measure, the probability associated with smaller and smaller intervals of Ω (around unity) at earlier and earlier times need not become arbitrarily small because of the singularity at $\Omega = 1$.

8

Indeed, this measure is constructed in precisely such a way that the probability associated with a given range of Ω_0 is preserved as the system evolves. We should not therefore be surprised to find $\Omega_0 \simeq 1$ at the present epoch even in the absence of inflation, so we do not need inflation to "explain" this value. In this sense, *there is no flatness problem* in a purely classical cosmological model.

We realise that many of the issues we have discussed remain controversial. We accept, for example, that Jaynes' principle may be the last word in the theory of prior assignment based on minimal information. Nevertheless, inferences based only on vague prescriptions of uniform priors have no place in physics or cosmology. Consistent inverse reasoning *requires* the assignment of a prior according to some objective rules; failure to do this replaces bona fide inductive logic with mere superstition.

9

REFERENCES

(1) P. Coles & G.F.R. Ellis, 1994. *Nature,* **370**, 609–615.

(2) A.H. Guth, 1981. *Phys. Rev. D.,* **23**, 347–356.

(3) A. Tarantola & B.J. Valette, 1982. *J. Geophys.,* **50**, 159–170.

(4) R.T. Cox, 1946. *Am. J. Phys.,* **14**, 1–13.

(5) A.J.M. Garrett & P. Coles, 1993. *Comments on Astrophys.,* **17**, 23–47.

(6) T.J. Loredo, 1990. In *Maximum Entropy and Bayesian Methods,* ed. P.F. Fougère, pp. 81–142, Kluwer, Dordrecht.

(7) E.T. Jaynes, 1968. *IEEE Transactions on Systems Science and Cybernetics,* **SSC–4**, 227–241.

(8) G. Evrard, 1995, *Physics Letters,* **A201**, 95–102.

(9) H. Jeffreys, 1939. *Theory of Probability,* Clarendon Press, Oxford.