

An extension of Friedmann–Robertson–Walker theory beyond big bang

Joachim Schröter

Department of Theoretical Physics, University of Paderborn, D-33095
Paderborn, Germany
J@schroe.de

Abstract

Starting from the classic Friedmann–Robertson–Walker theory with big bang it is shown that the solutions of the field equations can be extended to negative times. Choosing a new cosmic time scale instead of proper time one achieves complete differentiability of the scale factor and of suitable thermodynamic quantities equivalent to pressure and energy density. Then, the singularity of big bang manifests itself only by the vanishing of the scale factor at time zero. Moreover, all solutions of the field equations are defined for all times from $-\infty$ to $+\infty$. In a separate section, the horizon structure of the extended theory is studied. Some weak assumptions guarantee that there are no horizons. Hence, the horizon problem in a strict sense disappears. An intensive discussion of the results is given at the end of the paper.

1 Introduction

1.1: Customarily, the term Friedmann–Robertson–Walker theory (abbreviated FRW) denotes that branch of General Relativity that deals with homogeneous and isotropic space-times and thus is of interest for a rough description of the universe. FRW is the starting point of the following considerations. For the sake of definiteness let us briefly write down the basic features of FRW as they are used later on.

FRW1: The space-time manifold M^* of FRW is defined by

$$M^* = N_k \times I^*, \quad k = 0, \pm 1, \quad (1.1)$$

where $N_0 = \mathbb{R}^3$, N_1 is the 3-sphere S^3 and N_{-1} is the 3-hyperboloid H^3 . Moreover, I^* is either a finite or an infinite interval of the reals. (Other possible realizations of N_k are not taken into account here!). Matter is at rest in N_k .

FRW2: The metric g^* of FRW is given by

$$g^* = K^2 h_k - dt \otimes dt, \quad k = 0, \pm 1. \quad (1.2)$$

The scale factor K depends only on $t \in I^*$, and h_k is the Riemannian metric on N_k , which is induced by the Euclidean metric of the imbedding space \mathbb{R}^4 of N_k . The time t is the proper time of the matter fixed in N_k .

FRW3: The field equations of FRW are the Einstein equations, the balance of energy and momentum, and equations of continuity. These equations determine the fields K , p^* and e^* , where K is the scale factor, p^* the total pressure and e^* the total energy density of the considered system. They only depend on time t . The functions p^* and e^* are related to the thermodynamic state of the considered system by constitutive equations.

The Einstein equations, including a cosmological constant Λ_0 , can be brought into the following form:

$$2 \frac{\ddot{K}}{K} + \frac{\dot{K}^2}{K^2} + \frac{k}{K^2} - \Lambda_0 = -\kappa_0 p^*, \quad (1.3)$$

$$3 \frac{\dot{K}^2}{K^2} + 3 \frac{k}{K^2} - \Lambda_0 = \kappa_0 e^*, \quad (1.4)$$

$$6 \frac{\ddot{K}}{K} = -\kappa_0 \left(3p^* + e^* - \frac{2}{\kappa_0} \Lambda_0 \right). \quad (1.5)$$

The balance of energy and momentum reads:

$$3\frac{\dot{K}}{K}(p^* + e^*) + \dot{e}^* = 0. \quad (1.6)$$

The equations of continuity are not written down here because they depend specifically on the considered material system. (As is generally known, (1.3) to (1.6) are not independent.)

1.2: A short inspection of FRW 1 to 3 shows that with respect to geometry, FRW is almost fixed. The only variable term is the scale factor K in the metric g^* . It is an unknown function of time t . The enormous applicability of FRW is due to the thermodynamic side of the theory. It has its origin in the constitutive equations for p^* and e^* . These equations, together with the cosmological constant Λ_0 , which has to be regarded as a constitutive element of FRW, define the various models of the theory.

The literature about FRW is “almost infinite”. Therefore I confine myself to the citing of the basic papers by Friedmann, Robertson and Walker [1–5] and to some monographs from which I derived special advantage [6, 7, 9–12]. Many others could be cited as well.

1.3: In what follows, attention is mainly focused on that class of models of FRW, which exhibit the phenomenon called big bang.

With respect to the scale factor K , this class is characterized by the following four conditions:

B1: $K(t) > 0$ for $t \in I^*$.

B2: There is a time $t_1 \in I^*$ for which $\dot{K}(t_1) > 0$.

B3: $I^* =]0, T[$ where T is finite or infinite.

B4: $\lim_{t \rightarrow 0} K(t) = 0$.

Condition B4 together with its implications relating to p^* and e^* constitutes what usually is called big bang. (For the motivation of B1 to B4 cf. e.g., the cited literature.)

In order to have a simple and convenient notation I use in the following the term FRW in a somewhat restricted sense: it is characterized by FRW1 to FRW3 together with the conditions B1 to B4. This means that in this paper we are (almost) always concerned within the frame of FRW1 to 3 with cosmological models exhibiting big bang. Only in Section 2.6, some remarks are added about other models, which are connected to an extension of FRW.

1.4: It is well known that big bang is not an event in the sense of General Relativity, i.e., there is no point $x \in M^*$ which corresponds to a natural phenomenon we call big bang. Nevertheless, the question is legitimate if it is possible to change the theory in such a way that big bang is describable as an event, i.e., as a point of some Lorentz manifold. Clearly, as big bang is theoretically understood to be a singularity there is a problem to what extend it can be regularized. By these remarks the problem we are dealing with in this paper is briefly outlined.

2 The extended Friedmann–Robertson–Walker (EFRW) theory

2.1: In a first step of extending FRW, a proposition is formulated and proved which is the key towards EFRW. For this purpose, we need some

Notation. 1. Let K be the scale factor of FRW, which is defined on $I^* =]0, T[$.

Then define R_{\pm} by

$$R_{\pm}(t) = \begin{cases} K(t), & t \in I^*, \\ \pm K(-t), & t \in -I^*. \end{cases} \quad (2.1)$$

Moreover let the functions p and e be defined by

$$p(t) = p^*(|t|), \quad e(t) = e^*(|t|), \quad (2.2)$$

where $t \in I^* \cup -I^*$. (The notation implies that t is the same parameter as in FRW!)

2. Since in what follows we are mainly concerned with R_- , notation is simplified by omitting the signs $+$ or $-$. If necessary, we write $R = R_+$ or $R = R_-$.

Now the basic theorem can be stated thus.

Proposition. Let K, p^*, e^* be a solution of (1.3) to (1.6). Then both $R = R_+, p, e$ and $R = R_-, p, e$ are solutions of the equations

$$2\frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{k}{R^2} - \Lambda_0 = -\kappa_0 p, \quad (2.3)$$

$$3\frac{\dot{R}^2}{R^2} + 3\frac{k}{R^2} - \Lambda_0 = \kappa_0 e, \tag{2.4}$$

$$6\ddot{R} + \kappa_0 R \left(3p + e - \frac{2}{\kappa_0} \Lambda_0 \right) = 0, \tag{2.5}$$

$$3\frac{\dot{R}}{R}(p + e) + \dot{e} = 0 \tag{2.6}$$

for all $t \in I^* \cup -I^*$. Equations (2.3) to (2.6) have the same interdependences as (1.3) to (1.6).

Proof. Since for $t \in I^*$ the relations $R(t) = K(t), p(t) = p^*(t)$ and $e(t) = e^*(t)$ hold, (2.3) to (2.6) are identical with (1.3) to (1.6). If $t \in -I^*$, one finds that

$$R(t) = \pm K(-t), \quad \dot{R}(t) = \mp \dot{K}(-t), \quad \ddot{R}(t) = \pm \ddot{K}(-t) \tag{2.7}$$

and

$$p(t) = p^*(-t), \quad e(t) = e^*(-t), \quad \dot{e}(t) = -\dot{e}^*(-t). \tag{2.8}$$

Now inserting (2.7) and (2.8) into (2.3) to (2.6) one again finds by a simple calculation that (1.3) to (1.6) hold. Thus, the proposition is seen to be true.

Some special cases of the proposition can be found in the literature.

Since the domain of a solution $R = R_{\pm}, p$ and e of (2.3) to (2.6) is the set $I^* \cup -I^*$, the functions R, p, e are not defined for $t = 0$. Thus the problem sketched in Section 1.4 now reads:

1. Is it possible to define the triple (R, p, e) or some equivalent also for $t = 0$?
2. Are the considered functions, if definable for $t = 0$, also continuous or even differentiable at this point?

This problem will be treated for $R = R_-$ in Sections 2.2 to 2.5. Some remarks concerning R_+ are added in Section 2.6.

2.2: In this subsection, a function R is considered which is the first component of a solution (R, p, e) of (2.3) to (2.6). More specifically, we are interested in the behaviour of R in a neighbourhood of $t = 0$. From the definition of R we arrive at the following two

Consequences. 1. *Since by assumption only such models of FRW are taken into account for which $\lim_{t \rightarrow 0^+} K(t) = 0$ we find that for each R the relation*

$$\lim_{t \rightarrow 0} R(t) = 0 \quad (2.9)$$

holds for arbitrary limits. Hence, if we define $R(0) = 0$, the function R has the domain $I =]-T, T[$ and is continuous.

2. *By definition of R (with $R = R_-$) the relation $\dot{R}(-t) = \dot{R}(t)$ holds. Therefore from (2.4), one concludes that*

$$\lim_{t \rightarrow 0} \dot{R}(t) := a_0 < \infty \quad (2.10)$$

exactly if

$$\lim_{t \rightarrow 0} e(t)R(t)^2 := b_0 < \infty. \quad (2.11)$$

3. *By (2.1), (2.2) and (2.9) the phenomenon of a big bang in FRW is replaced by a big crunch bang in EFRW. This means, if (2.11) is valid, the function R is continuously differentiable in the whole domain I of R . However, as we shall see, condition (2.11) is not fulfilled by realistic constitutive equations for e .*

2.3: The question how the thermodynamic quantities p and e behave for $t \rightarrow 0$ can only be answered if some assumptions about the constitution of the universe for $t \rightarrow 0$ are imposed. The generally accepted supposition is the following:

The universe is radiation dominated in a neighbourhood of $t = 0$ in the sense that the influence of ponderable matter is asymptotically negligible if $t \rightarrow 0$. In other words, the model of a radiation dominated universe gives an asymptotically correct description of the real situation in the vicinity of $t = 0$

This assumption has the following

Consequences. 1. *Let η denote the density of photons. Then $e = \eta$ and $p = \frac{1}{3}\eta$, so that the solution of (2.6) reads*

$$\eta = AR^{-4}, \quad A = \eta(t_0)R(t_0)^4. \quad (2.12)$$

For the sake of simplicity let us assume that $\Lambda_0 = 0$. Then the solution of (2.4) and (2.12) is given by

$$R(t)^2 = \left(\frac{4}{3}\kappa_0 A\right)^{\frac{1}{2}} |t| - kt^2, \quad k = 0, \pm 1 \tag{2.13}$$

and

$$R(t) = \pm(R(t)^2)^{\frac{1}{2}}, \quad \text{if } \pm t \geq 0. \tag{2.14}$$

The result is well known for $t > 0$.

2. If one considers the specific photonic volume $\omega = \eta^{-1}$ instead of η , then by (2.12), one finds that $\omega = A^{-1}R^4$. Hence ω is continuous at $t = 0$ but not differentiable because R is not differentiable. Equations (2.3) and (2.4) now take the form

$$(2R\ddot{R} + \dot{R}^2 + k - \Lambda_0 R^2)R^2 = -\frac{1}{3}\kappa_0 A, \tag{2.15}$$

$$(3\dot{R}^2 + 3k - \Lambda_0 R^2)R^2 = \kappa_0 A. \tag{2.16}$$

3. Equations (2.15) and (2.16) are used to determine the asymptotic behaviour of \dot{R} and \ddot{R} . Since, by supposition the relation (2.9) holds, one concludes from (2.16) that

$$\lim_{t \rightarrow 0} \dot{R}^2 R^2 = \frac{\kappa_0}{3} A > 0.$$

Hence,

$$\lim_{t \rightarrow 0} \dot{R}|R| = \left(\frac{\kappa_0}{3}\right)^{\frac{1}{2}} A^{\frac{1}{2}} \tag{2.17}$$

or

$$\dot{R} = \mathcal{O}(|R|^{-1}) \quad \text{for } t \rightarrow 0. \tag{2.18}$$

The derivative \dot{R} diverges like $|R|^{-1}$. From (2.15), we see that

$$\lim_{t \rightarrow 0} R^3 \ddot{R} = -\frac{1}{3}\kappa_0 A^{-1} < 0 \tag{2.19}$$

or

$$|\ddot{R}| = \mathcal{O}(|R|^{-3}) \quad \text{for } t \rightarrow 0. \quad (2.20)$$

Hence, the second derivative \ddot{R} diverges like $|R|^{-3}$.

2.4: Equations (2.17) to (2.20) suggest asking whether it is possible to obtain differentiability of R by a transformation of the time parameter t .

2.4.1: This is indeed possible as the following considerations show.

Let f be a bijective function, which is two times continuously differentiable and satisfies the relation $t = f(\tau) = -f(-\tau)$. Now define S by $S(\tau) = R(f(\tau))$. Then it follows from (2.7) that

$$S(\tau) = -S(-\tau), \quad (2.21)$$

$$\dot{S}(\tau) = \dot{S}(-\tau), \quad \ddot{S}(\tau) = -\ddot{S}(-\tau) \quad \text{for } \tau \neq 0. \quad (2.22)$$

Therefore, equation $S(0) = 0$ is a necessary condition. Moreover, if S is two times continuously differentiable in I , it is also necessary that $\dot{S}(0) = 0$ holds. Now let us consider sufficient conditions for S to be differentiable. By definition of S we find for $\tau \neq 0$ that

$$\dot{S}(\tau) = \dot{R}(f(\tau))\dot{f}(\tau), \quad (2.23)$$

$$\ddot{S}(\tau) = \ddot{R}(f(\tau))(\dot{f}(\tau))^2 + \dot{R}(f(\tau))\ddot{f}(\tau). \quad (2.24)$$

From (2.18) and (2.23), one concludes that $\dot{S}(0)$ exists if

$$\dot{f}(\tau) = \mathcal{O}(|S(\tau)|) \quad \text{for } \tau \rightarrow 0. \quad (2.25)$$

A similar result holds for the second derivative. From (2.20) to (2.24) one derives that $\ddot{S}(0)$ exists and that $\ddot{S}(0) = 0$, if

$$\dot{f}(\tau) = o(|S(\tau)|^{\frac{3}{2}}), \quad \ddot{f}(\tau) = o(|S(\tau)|), \quad \text{for } \tau \rightarrow 0. \quad (2.26)$$

2.4.2: These general results now are illustrated by a significant example. Let us consider the radiation dominated era with $\Lambda_0 = 0$. Then according

to (2.13) and (2.14), we obtain with $B := \frac{4}{3}\kappa_0 A$ that

$$R(t) = \pm(B|t| - kt^2)^{\frac{1}{2}}, \quad \pm t \geq 0. \tag{2.27}$$

Now define $t = f(\tau) := \alpha\tau^n$ where n is an odd positive integer and $\alpha > 0$. Inserting $f(\tau)$ into (2.23) and (2.24) one obtains

$$\dot{S}(\tau) = |\tau|^{\frac{n}{2}-1}F_1(\tau), \quad \ddot{S}(\tau) = \pm|\tau|^{\frac{n}{2}-2}F_2(\tau), \tag{2.28}$$

where F_1 and F_2 are continuously differentiable functions and where

$$F_1(0) = \frac{n}{2}\alpha^{\frac{1}{2}}B^{\frac{1}{2}}, \quad F_2(0) = \frac{n}{4}(2n - 3)\alpha^{\frac{1}{2}}B^{\frac{1}{2}}. \tag{2.29}$$

From (2.28), we draw the conclusion that the transformed scale factor S is two times continuously differentiable if $n \geq 5$ and that the equations $S(0) = 0, \dot{S}(0) = 0$ and $\ddot{S}(0) = 0$ hold.

2.4.3: The example $f(\tau) = \alpha\tau^n$ of a time transformation shows the special feature that the proper time t is totally replaced by time τ . But proper time t has proved its worth far from $t = 0$ or $\tau = 0$, respectively. Therefore, one should define f for large t by $t = f(\tau) := \tau$. Qualitatively the function f should have the form as illustrated in figure 1.

In case that T is finite it is useful to change the function f defined on $[-T, T]$ in a neighbourhood of $-T$ and of T in the same way as it was done in a neighbourhood of $\tau = 0$. Then qualitatively f looks as in figure 2.

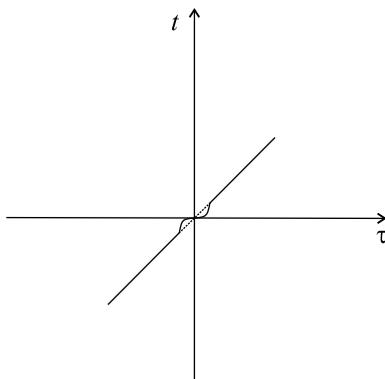


Figure 1: $t = f(\tau)$ for infinite T .

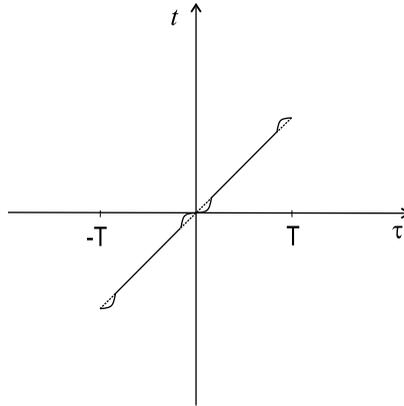


Figure 2: $t = f(\tau)$ for finite T .

In this case, f satisfies the conditions

$$\dot{f}(\pm T) = 0, \quad \ddot{f}(\pm T) = 0. \quad (2.30)$$

If f is defined this way, the solutions of the field equations defined in the intervals $[(2z - 1)T, (2z + 1)T], z \in \mathbb{Z}$ can be connected so that an infinitely periodic scale factor results.

2.4.4: Special attention with respect to differentiability and continuity has to be paid to the constitutive quantities p and e . As already remarked in Section 2.3 (Consequence 2.), a transformation of the time scale $t = f(\tau)$ is not sufficient for removing the divergence of p and e at $t = 0$ or $\tau = 0$. The pressure p and the energy density e are not suitable quantities to describe big crunch bang smoothly! Which quantities are suitable? In order to find an answer let us consider $\bar{p}(\tau) = p(t)$, $\bar{e}(\tau) = e(t)$, $t = f(\tau)$, and in addition the quantities $\psi = \bar{p}^{-1}$ and $\omega = \bar{e}^{-1}$. Moreover, since S is bijective in a neighbourhood of $\tau = 0$ we obtain $\psi(\tau) = \hat{\psi}(S(\tau))$, $\omega(\tau) = \hat{\omega}(S(\tau))$, and arrive at the following

Result. *The constitutive quantities ψ and ω are suitable for a smooth description of big crunch bang if the derivatives of $\hat{\psi}$ and $\hat{\omega}$ up to the second order are continuous in a neighbourhood of $\tau = 0$.*

2.4.5: If the time transformation f is different from id , i.e., if $t = \tau$ does not hold throughout, the fields (2.3) to (2.6) have to be modified. This can be achieved most easily by inserting $R(t) = S(\tau)$, $p(t) = \bar{p}(\tau)$ and $e(t) = \bar{e}(\tau)$ together with $\tau = f^{-1}(t)$ into (2.3) to (2.6). The result is given in Section 2.5 formula (2.33).

2.5: The heuristic considerations of Sections 2.1 to 2.4 now lead to a compact description of EFRW.

2.5.1: The basic properties of EFRW are condensed in the following four conditions:

EFRW 1: The manifold of events M is given by $M = N_k \times I$, $k = 0, \pm 1$, where $N_0 = \mathbb{R}^3$, N_1 is the 3-sphere and N_{-1} is the 3-hyperboloid. Moreover, $I =]-T, T[$. There is a global chart on I the coordinate function of which is the identity id . The time parameter is denoted τ and is called the (new) cosmic time scale.

EFRW 2: 1. The metric g is defined on M . It has the form

$$g = S^2 h_k - \dot{f}^2 d\tau \otimes d\tau. \quad (2.31)$$

2. The tensor h_k is the Riemannian metric on N_k , which is induced by the Euclidean metric of the imbedding space \mathbb{R}^4 of N_k .

3. The scale factor S depends only on τ . It is of class C^r , $r \geq 2$.

4. The time parameter τ and the proper time t of the matter fixed in N_k are related by an equation $t = f(\tau)$, where f is an increasing function of class C^l , $l \geq 3$, which is given as a constitutive element satisfying the condition

$$f(-\tau) = -f(\tau). \quad (2.32)$$

Moreover, the inverse function f^{-1} is also of class C^l for $t \neq 0$.

EFRW 3: 1. The field equations are (2.3) to (2.6), where the time t is transformed by $t = f(\tau)$. The constitutive quantities \bar{p} and \bar{e} defined by $\bar{p}(\tau) = p(f(\tau))$, $\bar{e}(\tau) = e(f(\tau))$ can be replaced by suitable quantities ψ, ω such that the field equations contain only the functions S, ψ, ω , which are defined for each $\tau \in I$. If $\bar{p} \neq 0$ and $\bar{e} \neq 0$ the quantities $\psi = \bar{p}^{-1}$ and $\omega = \bar{e}^{-1}$ are suitable. In this case, the field equations read:

$$\begin{aligned} \psi(2Sq\ddot{S} + q\dot{S}^2 - 2S\dot{q}\dot{S} + kq^3 - \Lambda_0q^3S^2) &= -\kappa_0q^3S^2, \\ \omega(3\dot{S}^2 + 3kq^2 - \Lambda_0q^2S^2) &= \kappa_0q^2S^2, \\ 6q\omega\psi\ddot{S} &= -\kappa_0q^3S(3\omega + \psi) + 6\psi\omega\dot{q}\dot{S} + 2\Lambda_0q^3S\omega\psi, \\ 3\omega^2(\omega + \psi)\dot{S} - \psi\omega S\dot{\omega} &= 0, \end{aligned} \quad (2.33)$$

where $q = \dot{f}$.

2. The functions S, ψ, ω satisfy the following relations:

$$\begin{aligned} S(\tau) &= -S(-\tau), & \psi(\tau) &= \psi(-\tau), & \omega(\tau) &= \omega(-\tau), & \text{for all } \tau \in I, \\ S(\tau) &> 0, & \text{for } \tau > 0 & \text{ and } \dot{S}(0) = 0. \end{aligned} \quad (2.34)$$

Because of $\dot{R}(0) \neq 0$ and (2.23) it follows from (2.34) that $\dot{f}(0) = 0$.

EFRW 4: Outside a time interval $[-\epsilon, \epsilon]$, $\epsilon > 0$, the formulation of EFRW in terms of t, R, p, e is equivalent to that in terms of τ, S, ψ, ω . In this case, t is given by two charts on I , one for $t > 0$ the other for $t < 0$ (cf. Section 2.5.3.1).

2.5.2.: The strategy for solving (2.33) and (2.34) is extremely simple. It can be achieved in four steps.

1. Let the constitutive equations for p^* and e^* be given. Then, first of all one solves (1.3) to (1.6) from FRW with the initial condition $\lim_{t \rightarrow 0} K(t) = 0$.
2. If K, p^* and e^* are known one defines the functions R, p and e by (2.1) and (2.2). They satisfy (2.3) and (2.6).
3. By a time transformation $t = f(\tau)$ and the choice of suitable quantities, e.g., $\psi = p^{-1}, \omega = e^{-1}$ a solution S, ψ, ω of (2.33) and (2.34) is obtained.
4. If one is only interested in times unequal zero the two descriptions in terms of t, R, p, e and in terms of τ, S, ψ, ω are equivalent. For practical purposes the time scale t is generally more comfortable than the scale τ .

2.5.3: In this subsection, some remarks are added to illustrate the results of EFRW.

2.5.3.1: Since $R(0) = 0$, the metric g written down in coordinates with time t (i.e., g of the form (2)) is singular at $t = 0$. This means $g = dt \otimes dt$ and $\det g = 0$. After the time transformation $t = f(\tau)$ one obtains $dt = \dot{f}(\tau)d\tau$ so that $dt = 0$ for $\tau = 0$. Hence $g = 0$ for $\tau = 0$, because $S(0) = 0$ and $\dot{f}(0) = 0$. These properties show that the time scales t and τ are not compatible at $t = 0$ or $\tau = 0$, respectively. What is the right scale?

On mathematical grounds, τ is preferable because $S(\tau) = R(f(\tau))$ is differentiable for all $\tau \in I$, whereas $R(t)$ is not differentiable at $t = 0$. Expressed in geometrical terms this means, the identity function id is a global time chart on I , whereas the time $t = f(\tau)$ is not given by a global time chart because f^{-1} is not differentiable at $t = 0$, but only continuous throughout. Thus, if we introduce the charts f_{\pm} by $f_{\pm}(\tau) = f(\tau)$ for $\pm\tau > 0$ time t is seen to be given by two charts C^l -compatible ($l \geq 3$) with τ . Both these

charts can sometimes be used for convenient calculating. The change of the cosmic time scale from t to τ is one of the essential features of EFRW. Nevertheless, there is one special point: $g = 0$ for $\tau = 0$. We shall come back to it in Section 4.1.4. Physical arguments in favour of τ are given in Section 4.2.

2.5.3.2: As already pointed out in Section 1.2, the constitutive equations for p^* and e^* and the possible values of Λ_0 define the different models of FRW and consequently the models of EFRW. This implies that all results obtained in FRW (in the sense of Section 1.3) are also valid in EFRW if the asymptotic behaviour of R (cf. (2.17) to (2.20)) is guaranteed. This implies $\lim_{t \rightarrow 0^+} K(t) = 0$. But it is well known that there are constitutive scenarios p^*, e^*, Λ_0 such that K does not satisfy this condition (cf. e.g., [7]).

As usual, if cosmological problems are treated within the frame of FRW the constitutive equations cannot be specified by only one function of the thermodynamic state for the whole time interval I^* , rather one uses different constitutive equations for different subintervals of I^* . EFRW exhibits the same features. But one can even go a step further. The constitutive scenario for negative times τ is in EFRW totally determined by the scenario for positive τ . If there are reasons to impose other constitutive relations for $\tau < 0$ this can be done very easily. One only has to set up two models of FRW and glue them together according to the rules developed in the previous sections. Up to now, no such reasons are known!

2.5.3.3: It was already mentioned at the end of Section 2.4.3 that models of FRW with finite-time intervals I^* can be glued together if the time scales t and τ are related by a function f as qualitatively illustrated in figure 2. This means that each model of FRW with finite I^* generates a periodic model of EFRW, i.e., a model with infinite I . As long as there is no reason for considering only a finite part of a periodic model all models in EFRW are defined for all $\tau \in I = \mathbb{R}$.

2.5.3.4: Summing up, EFRW is characterized by the following features: the new cosmic time scale τ , the use of suitable constitutive quantities, e.g., ψ, ω , and the infinite time interval $I = \mathbb{R}$.

2.6: The considerations of Sections 2.2 to 2.5. referred only to $R := R_-$ and $S = R(f)$. But the results there obtained are almost all valid also for R_+ and $S_+ = R_+(f)$. One of the differences between $R := R_-$ and R_+ , and S and S_+ , respectively, is that R and S are monotone functions in an interval containing big crunch bang, hence invertible, whereas R_+ and S_+ are not. In the next section, this property of R will be exploited.

Nevertheless, R_+ comes into play if in FRW scenarios without big bang are considered. A nontrivial example of this kind was given by Lessner (cf. [13]). I will come back to these aspects of EFRW in a subsequent paper.

3 Horizons

3.1 General remarks

3.1.1: The concept of horizon is well established in FRW (cf. [6–12]). In this subsection it will be introduced also in EFRW. More precisely, we are looking for a common concept in FRW and EFRW without altering the results in FRW. For this purpose, we need to fix some

Notation. 1. Let the manifold of FRW be $M^* = N_k \times I^*$ and let that of EFRW be $M = N_k \times I$. Then in both these cases the time coordinate is chosen to be proper time t . This choice is possible because we want to study integrals and for these purposes t is more comfortable than time τ . Moreover, if N_k has to be coordinatized we take the triple $(\chi, \vartheta, \varphi)$ where χ is the radial distance, and ϑ, φ , as usual, are the polar angles.

2. The scale factor, the pressure and the energy density are denoted in FRW and in EFRW by R, p, e .

3. In EFRW let $t^* := \inf I = -\infty$ and $t^\sharp := \sup I = \infty$ and in FRW let $t^* := \inf I^* = 0$ whereas $t^\sharp := \sup I^*$ is finite or infinite.

4. The function J is defined by

$$J(t, t') = \int_{t'}^t |R(s)|^{-1} ds \quad (3.1)$$

for each pair t, t' for which the integral exists.

Then, we obtain the following

Consequences. 1. If $t, t' > 0$ or if $t, t' < 0$ the integrand in (3.1) is bounded. Hence, $J(t, t')$ exists and $J(t, t') = J(-t', -t)$.

2. In order to evaluate $J(t, t')$ for $t, t' > 0$ or for $t, t' < 0$, one need not solve the field equations. It suffices to know the constitutive equation $e = \hat{e}(R)$ and the two values $R(t)$ and $R(t')$.

The proof runs as follows. First, let $\dot{R}(s) > 0$ for $t' \leq s \leq t$. Then the equation, $\sigma = R(s)$ is solvable for s so that σ can be used as integration

variable. Therefore

$$J(t, t') = \int_{R(t')}^{R(t)} |\sigma|^{-1} \dot{R}^{-1} d\sigma. \tag{3.2}$$

Inserting (2.4) into (3.2), one obtains

$$J(t, t') = 3^{\frac{1}{2}} \int_{R(t')}^{R(t)} |\sigma|^{-1} (\kappa_0 \hat{e}(\sigma) \sigma^2 + \Lambda_0 \sigma^2 - 3k)^{-\frac{1}{2}} d\sigma. \tag{3.3}$$

If \dot{R} changes sign in $[t', t]$ then this interval is the union of subintervals for which (3.3) holds, and $J(t, t')$ is a sum, which adds up to (3.3).

With the help of (3.3) now it is easy to check if $J(t, t')$ exists for $t' < 0 < t$.

Proposition. *If matter is asymptotically radiation dominated for $t, t' \rightarrow 0$, then $J(t, t')$ exists for $t' < 0 < t$.*

Proof. It suffices to prove $J(t, 0) < \infty$ for $t > 0$. By supposition $\hat{e}(\sigma) = B\sigma^{-4}$ with $B > 0$, and $0 \leq \sigma \leq R(\bar{t})$ for some $\bar{t} > 0$. Now let

$$F(\sigma) := \kappa_0 B + \Lambda_0 \sigma^4 - k\sigma^2.$$

Then there is $m > 0$ such that $F(\sigma) > m$ for sufficiently small σ , i.e., for sufficiently small \bar{t} . The integrand in $J(\bar{t}, 0)$ according to (3.3) is $F(\sigma)^{-\frac{1}{2}}$. Hence,

$$J(\bar{t}, 0) \leq 3^{\frac{1}{2}} m^{-\frac{1}{2}} R(\bar{t}) < \infty. \tag{3.4}$$

Because of $J(t, 0) = J(t, \bar{t}) + J(\bar{t}, 0)$ the propositions holds.

3.1.2 Particle horizons

3.1.2.1: Since ponderable matter is assumed to be at rest in N_k a particle can be identified with its position in N_k . Consequently, the trajectory γ_a of a particle $a \in N_k$ is given by $\gamma_a(t) = (a, t)$ with $t \in I$ or I^* respectively, and the worldline of a is $W_a := \text{ran } \gamma_a$. Then the concept of a particle horizon is defined as follows:

Definition. The particle horizon $\mathcal{H}(b, t)$ of an observer b at time t is the boundary between the set of particles $\mathcal{K}(b, t)$, which can send signals arriving at b up to time t , and the set of particles which can not.

If $\mathcal{H}(b, t) = \emptyset$, it is usual to say that b does not have a particle horizon at time t .

3.1.2.2: The homogeneity and isotropy of the manifolds considered in EFRW and FRW allow us to formulate a simple criterion for the existence of a particle horizon as follows:

Proposition. 1. Case $k = 0, -1$. The particle horizon $\mathcal{H}(b, t)$ of an observer b at time $t > t^*$ exists exactly if $J(t, t^*) < \infty$.

2. Case $k = 1$. The particle horizon $\mathcal{H}(b, t)$ of an observer b at time $t > t^*$ exists exactly if $J(t, t^*) < \pi$.

3. In both these cases the particle horizon is given by

$$\mathcal{H}(b, t) = \partial\mathcal{K}(b, t)$$

and

$$\mathcal{K}(b, t) = \{(\chi, \vartheta, \varphi) : 0 \leq \vartheta \leq \pi, \quad 0 \leq \chi \leq J(t, t^*), \quad \varphi \in [0, 2\pi)\}. \quad (3.5)$$

The *proof* can be extracted from the literature, especially from [7, 8, 12].

Immediately from the proposition one obtains the following

Consequences. 1. If $\mathcal{H}(b, t) = \emptyset$ and if $t' > t$ then also $\mathcal{H}(b, t') = \emptyset$, because $J(t', t^*) \geq J(t, t^*)$.

2. If $\mathcal{H}(b, t) \neq \emptyset$ and if $t'' < t$ then also $\mathcal{H}(b, t'') \neq \emptyset$, because $J(t'', t^*) < J(t, t^*)$.

The physical significance of the concept of a horizon is underlined by the following:

Proposition. $b_2 \in \mathcal{K}(b_1, t)$ if and only if $b_1 \in \mathcal{K}(b_2, t)$.

The *proof* can be read off from (3.5) if one takes into account that if $b_1 = (\chi_1, \vartheta_1, \varphi_1)$ and $b_2 = (\chi_2, \vartheta_2, \varphi_2)$ in their respective coordinate systems, we have $\chi_1 = \chi_2$, because χ_1 is the distance between b_2 and b_1 in the Riemannian space (N_k, h_k) , $k = 0, \pm 1$, and χ_2 is the distance between b_1 and b_2 .

The proposition can be physically interpreted: let b be a particle. Then b is causally affected by all particles a out of the horizon of b at time t and its interior, and b affects causally up to time t all particles out of its horizon and its interior (at t). Thus, if there are no horizons all particles are in causal contact with each other.

3.1.3 Event horizons

3.1.3.1: In this subsection, again the notation of Section 3.1.1 is used especially $t^\sharp := \sup I$ or $t^\sharp := \sup I^*$. Then the concept of an event horizon is given by the following:

Definition. *The event horizon $\mathcal{E}(b, t)$ of an observer b at time t is the boundary of the set $\mathcal{L}(b, t)$ of events $(a, t) \in N_k \times \{t\}$ from which signals reach b up to time t^\sharp and of the set of events $(a', t) \in N_k \times \{t\}$ from which signals cannot reach b .*

3.1.3.2.: For the existence of event horizons, a simple criterion also exists.

Proposition. 1. *Case $k = 0, -1$. The event horizon $\mathcal{E}(b, t)$ of an observer b at time $t < t^\sharp$ exists exactly if $J(t^\sharp, t) < \infty$.*

2. *Case $k = 1$. The event horizon $\mathcal{E}(b, t)$ of an observer b at time $t < t^\sharp$ exists exactly, if $J(t^\sharp, t) < \pi$.*

3. *In both these cases the event horizon is given by*

$$\mathcal{E}(b, t) = \partial\mathcal{L}(b, t)$$

and

$$\mathcal{L}(b, t) = \{(\chi, \vartheta, \varphi, t) : 0 \leq \vartheta \leq \pi, \quad 0 \leq \varphi \leq 2\pi, \quad \chi \leq J(t^\sharp, t)\}. \quad (3.6)$$

Again there are some immediate

Consequences. 1. *If $\mathcal{E}(b, t) = \emptyset$ and if $t' < t$ then also $\mathcal{E}(b, t') = \emptyset$, because $J(t^\sharp, t') \geq J(t^\sharp, t)$.*

2. *If $\mathcal{E}(b, t) \neq \emptyset$ and if $t'' > t$ then also $\mathcal{E}(b, t'') \neq \emptyset$, because $J(t^\sharp, t'') < J(t^\sharp, t)$.*

3.1.4: For a discussion of the horizon structure of FRW I refer to the cited literature (cf. e.g., [7, 10–12]). For $k = 0, -1$, the existence of particle horizons follows directly from the proposition of Section 3.1.1.

3.2 Horizons in EFRW

3.2.1: One of the characteristics of EFRW is the symmetry with respect to time. Again the notation of Section 3.1.1 is used. Then, we obtain the following

Proposition. *A particle horizon for an observer b at time t exists exactly if an event horizon for b exists at time $-t$.*

Proof. By definition of J the equation $J(t, t') = J(-t', -t)$ holds, thus also if $t' = -\infty$. Using the criteria of 3.1.2 and 3.1.3 the proposition is seen to hold.

Hence for a discussion of horizons in EFRW we need to calculate only one of the integrals $J(-t, -\infty)$ or $J(\infty, t)$. In Section 2.5.3.3, it was shown that the models of EFRW with changing signs of $\dot{R}(t)$ can be periodically extended to the interval $I = \mathbb{R}$ in a natural way. Hence all models of EFRW have the same time axis $I = \mathbb{R}$.

3.2.2 Periodic models

This case is the simplest one in EFRW. For each periodic model we obtain the following

Proposition. *There is neither a particle horizon nor an event horizon for any particle $b \in N_k, k = 0, \pm 1$ at any $t \in I = \mathbb{R}$.*

Proof. Let T be the length of a period, and let t be any time. Then there is a time t' such that $0 < t' - t < T$ and $J(t', t) > 0$. Hence, $J(t', t) = J(t' + nT, t + nT)$ and

$$J(\infty, t) \geq \sum_{n=0}^{\infty} J(t' + nT, t + nT) = \infty. \quad (3.7)$$

3.2.3 Nonperiodic models

3.2.3.1: In this section, models are considered for which $\dot{R}(t) > 0$ for all $t \in I = \mathbb{R}$.

Moreover, it is assumed that there is a time $t_1 > 0$ such that for $t > t_1$ space-time is asymptotically matter dominated and matter is dust. Therefore,

one obtains from (2.6) the constitutive equations:

$$p = 0, \quad e = A|R|^{-3} \tag{3.8}$$

with $A = e(t_0)|R(t_0)|^3$. Then $e(t) = e(-t)$.

As usual, it is supposed that the universe near big bang is radiation dominated.

Finally, in what follows we will study only such models for which $R(t) \rightarrow \infty$ holds if $t \rightarrow \infty$. (For the other cases cf. e.g., [7]).

These suppositions have two

Consequences. 1. For $t > t_1$ (3.3) with (3.8) reads:

$$J(\infty, t) = 3^{\frac{1}{2}} \int_{R(t)}^{\infty} \sigma^{-1} (\kappa_0 A \sigma^{-1} + \Lambda_0 \sigma^2 - 3k)^{-\frac{1}{2}} d\sigma. \tag{3.9}$$

2. If $\Lambda_0 < 0$, the bracket in (3.9) becomes negative for large σ . Therefore, $\Lambda_0 < 0$ is not compatible with the above assumptions so that we have to study only the cases $\Lambda_0 = 0$ and $\Lambda_0 > 0$.

3.2.3.2: Let $\Lambda_0 = 0$ and let us assume that the above suppositions are valid. Then necessarily $k = 0, -1$, and one obtains the following result.

Proposition. *There is neither a particle horizon nor an event horizon for any particle $b \in N_k, k = 0, -1$ at any time $t \in I = \mathbb{R}$.*

Proof. Let us consider $J(\infty, t')$ for a time $t' > t_1 > 0$ out of the dust era of matter. Then (3.9) can be applied and the variable σ in (3.9) obeys the inequality $\sigma \geq R(t') > 0$. Therefore $\kappa_0 A \sigma^{-1} + 3|k| \leq \kappa_0 A R(t')^{-1} + 3|k| =: N(t')^{-2}$, where $N(t') > 0$. Hence for $t' > t_1$, we have

$$J(\infty, t') \geq 3^{\frac{1}{2}} N(t') \int_{R(t')}^{\infty} \sigma^{-1} d\sigma = \infty. \tag{3.10}$$

For each other time t , the equation

$$J(\infty, t) = J(\infty, t') + J(t', t) = J(-t, -\infty) \tag{3.11}$$

holds. Then by the criteria of Sections 3.1.2.2 and 3.1.3.2, the proposition follows at once.

Remark. *This result has consequences for what is called the horizon problem. Strictly speaking, where there are no horizons there cannot be a horizon problem. At least, if there is a problem it does not make sense to call it “horizon” problem in this case.*

3.2.3.3: Now let $\Lambda_0 > 0$ and $k = 0, -1$. Then with the above suppositions, we obtain the following

Proposition. *For each particle $b \in N_k, k = 0, -1$, and each time $t \in I = \mathbb{R}$, there is both a particle horizon and an event horizon.*

Proof. It follows from (3.9) that for t' in the dust era

$$\begin{aligned} J(\infty, t') &= 3^{\frac{1}{2}} \int_{R(t')}^{\infty} \sigma^{-2} (\Lambda_0 + \kappa_0 A \sigma^{-3} + 3|k|\sigma^{-2})^{-\frac{1}{2}} d\sigma \\ &\leq 3^{\frac{1}{2}} \Lambda_0^{-\frac{1}{2}} \int_{R(t')}^{\infty} \sigma^{-2} d\sigma < \infty. \end{aligned} \tag{3.12}$$

Hence with (3.11), the proposition is seen to hold.

3.2.3.4: Let us finally consider the case $\Lambda_0 > 0$ and $k = 1$. Then one can show that the above conditions are only satisfiable if Λ_0 is restricted by the stronger condition $\Lambda_0 > \beta^{-2}$, where $\beta = \frac{1}{2}\kappa_0 A > 0$ (cf. e.g., [7]). But, in order to discuss the present case completely we had to specify the constitutive equations between matter domination and radiation domination, which is outside the scope of this paper. Thus, we confine ourselves to the simplest achievable result.

Proposition. *For each $\Lambda_0 > \beta^{-2}$, there is a time $t_2 \geq t_1 > 0$ such that for each particle $b \in N_1$ a particle horizon exists for $t' < -t_2$ and an event horizon for $t > t_2$.*

Proof. It follows from (3.9) that for each $\Lambda_0 > \beta^{-2}$ the inequality

$$J(\infty, t) < 3^{\frac{1}{2}} \int_{R(t)}^{\infty} h(\sigma)^{-\frac{1}{2}} d\sigma =: X(t) \tag{3.13}$$

holds with

$$\begin{aligned} h(\sigma) &= 2\beta\sigma + \beta^{-2}\sigma^4 - 3\sigma^2 \\ &= \beta^{-2}\sigma(\sigma + 2\beta)(\sigma - \beta)^2. \end{aligned} \tag{3.14}$$

Therefore,

$$X(t) = 3^{\frac{1}{2}}\beta \int_{R(t)}^{\infty} (\sigma^2 + 2\beta\sigma)^{-\frac{1}{2}} |\sigma - \beta|^{-1} d\sigma. \quad (3.15)$$

The last integral is elementarily solvable. It turns out that it is finite, hence

$$\lim_{t \rightarrow \infty} X(t) = 0.$$

Thus, there is a time t_2 such that $X(t_2) \leq \pi$.

3.2.4: The results of Sections 3.2.2 and 3.2.3 can be summarized as follows:

Each particle of a universe in EFRW is at each time in causal contact with all other particles if the universe is either periodic or aperiodic and $\Lambda_0 = 0$. This is an immediate consequence of the fact that there are no horizons in these universes. Some remarks on the so-called horizon problem can be found in Section 4.4.

4 Discussion of the results

4.1 How singular is big crunch bang in EFRW?

4.1.1: In order to get a sound answer to this question one has to say some words about the notion of a singularity as far as necessary for the purposes of this paper. The exact definition of a singularity in a space-time manifold can be found, e.g., in the classic monograph of Hawking and Ellis [11] and in some other texts (cf. [19–21]). Besides this exactly defined notion other, more intuitive concepts of a singularity are used. They are connected with terms such as indefiniteness, non-differentiability, divergence etc. In what follows, I call these intuitive concepts weak singularities.

4.1.2: Let us first consider the manifold M as it is introduced in Section 2.5.1.

By definition, $M = N_k \times I$, $k = 0, \pm 1$, where N_k does not depend on time. Hence, there is nothing special with M . (The same is the case for the manifold M^* of FRW.)

The kinematics of matter in M is given by the trajectories γ_a of particles $a \in N_k$, which are defined by $\gamma_a(\tau) = (a, \tau)$, $\tau \in I$. For each $a \in N_k$ the curve γ_a is a time-like affine geodesic, i.e., the particles $a \in N_k$ are freely falling. The set of all trajectories is a smooth congruence on M , which is inextensible with respect to the nonaffine parameter τ . There is no singularity in the weak sense.

4.1.3: Let us now consider a light-like signal between two points $p_1 = (a_1, \tau_1)$ and $p_2 = (a_2, \tau_2)$ of M , and let (χ, ϑ, ϕ) be defined as in Notation 1. of Section 3.1.1 such that a_1 is the origin of the coordinate system. Moreover, let be $a_2 = (\chi_2, \vartheta, \phi)$ and let σ be defined by

$$\sigma(\tau) = (\psi(\tau), \vartheta, \varphi, \tau), \quad (4.1)$$

where

$$\psi(\tau) = \int_{\tau_1}^{\tau} \frac{\dot{f}(\lambda)}{|S(\lambda)|} d\lambda = J(f(\tau), f(\tau_1)) \quad (4.2)$$

with S and f as in Section 2.5.1. Then the following result holds.

Proposition. 1. σ is a light-like curve between p_1 and p_2 if $\chi_2 = \psi(\tau_2)$

2. σ is defined for all $\tau_1, \tau_2 \in \mathbb{R}$ and is of class C^2 if matter is radiation dominated for $\tau \rightarrow 0$ and if the relations (2.26) hold.

3. σ is an affine geodesic, which is defined for all $\tau \in \mathbb{R}$. Hence, it is inextensible.

The *proof* is simple but lengthy, only the case $k = 1$ needs some care.

Since each light-like signal can be brought into the special form (4.1) we have obtained the complete null-geodesic structure. It is of class C^2 with respect to the nonaffine parameter τ . Hence, there is also no singularity in the weak sense.

4.1.4: Things change if one takes the metric g into account. As already remarked in Section 2.5.3.1 the metric becomes singular in the sense that $g = 0$, so that the inverse of g , i.e., the contravariant tensor g^\sharp , does not exist at $\tau = 0$. However, this property does not affect M as a manifold. Rather the measurements of proper length and proper times degenerate at $\tau = 0$, and this singularity is not removable within EFRW. But, in all models and applications of the theory the metric does not show up as a tensor. It suffices to know the functions S and \dot{f} in (2.31). Therefore, the singularity of g boils

down to the two relations $S(0) = 0$ and $\dot{f}(0) = 0$. The theoretical fact that $g = 0$ for $\tau = 0$, recedes to the background because it does not affect the practical work.

4.1.5: It is often believed that the most characteristic feature of big bang or big crunch bang is the divergence of the constitutive quantities \bar{p} and \bar{e} , which is caused by $S(\tau) \rightarrow 0$ for $\tau \rightarrow 0$. But this is not the case, this kind of singularity can be removed totally!

As we have seen in Sections 2.3 and 2.4.3 the quantities \bar{p} and \bar{e} are not suitable for a theoretical description of big crunch bang. Following the result of Section 2.4.2 the quantities $\psi = \bar{p}^{-1}$ and $\omega = \bar{e}^{-1}$ are suitable for this purpose. Introducing such “suitable quantities” instead of \bar{p} and \bar{e} looks like a dirty trick. But this is not the case as I will demonstrate by three arguments.

1. Let us assume that big crunch bang exists as a natural phenomenon. Then its existence is not touched by its theoretical description with the help of suitable quantities. The contrary is the case. An extraordinary phenomenon needs special treatment. The only criterion is that the latter is unobjectionable with respect to physics and mathematics.

2. If a singularity of a certain kind is thought to be essentially determined by the divergence of some physical quantities then in each field theory an arbitrary number of singularities can be constructed as follows. Let x be a point of a space-time, and let A_1, \dots, A_n be fields of the theory. Now introduce new fields by $B_j = (A_j - A_j(x))^{-1}$, $j = 1, \dots, n$, then B_j diverges at x . Hence, B_1, \dots, B_n are not suitable to describe the physical situation at x .

3. The “dirty” trick, i.e., introducing suitable quantities, was already used when FRW was derived from General Relativity specialized to homogeneous and isotropic spaces. There the primary quantity is scalar curvative \hat{C} . But it is more adequate to introduce the scale factor K by $\hat{C} = kK^{-2}$.

These arguments show that introducing suitable quantities to remove divergent terms is far from being a dirty trick.

4.1.6: Summing up the results of Sections 4.1.2 to 5, we see that in EFRW the kinematic and thermodynamic structures of matter as well as the structure of light-like geodesics are smooth and inextensible if one uses the new cosmic time τ . These properties of EFRW are very satisfactory. But, the question remains if the Lorentz manifold (M, g) is singularity-free in the sense of Hawking, Ellis et al. [11, 19–21] or not. I think it is not! The reason is that the concept of a singularity (in this strong sense) depends

on the use of affine (or generalized affine) parameters, and this has the consequence that proper time comes into play. Thus we are confronted with the same situation as in FRW. But, as we will see in Section 4.2, proper time loses its physical meaning in the vicinity of big crunch bang, because standard clocks indicating proper time are not constructible (cf. [14]). Hence the analysis of a singularity of (M, g) in the strong sense reveals only a mathematical property of EFRW, which is not physically relevant because it does not affect the physical properties of matter and light-like signals. It is nothing but attributing the term singular to a certain hyperplane. This guarantees the physical applicability of EFRW. The situation is similar to that described in Section 4.1.5. The choice of unsuitable quantities can produce singularities in a wide sense. In the present case, proper time is not suitable.

4.2 The clock problem in EFRW

In Section 2.4, we have seen that there are mathematical reasons to work with a new cosmic time scale τ instead of the time scale t which is the proper time of the particles of ponderable matter. In this section, I will complement the considerations of 2.4 by a physical argument in favour of τ .

Proper time was first introduced as a mathematical term. Later on the problem was studied whether it is possible to define proper time as that time which is indicated by a special class of clocks, the so-called standard clocks, which are constructible (in principle!) solely by geometrical means. Such clocks exist (cf. [14–16]). They indicate the proper time of a particle, which shows up in the construction of a standard clock. Then by experience we know that the time scales of standard clocks and of atomic clocks coincide at the present era of earth. The latter clocks are only constructible with the help of quantum mechanical means.

Within dense matter as in the universe near big crunch bang the geometric construction principles of standard clocks lose their reference to reality completely, because the freely falling test particles needed for their construction do not exist (cf. [14]). Proper, i.e., standard time then is nothing but a theoretical concept. This is a strong hint that it should be replaced by another time scale. This was done in Section 2.4 by introducing the new cosmic time scale τ . But, up to now it is an open problem whether there are constructive principles based on quantum theory for the clocks indicating time τ . Nevertheless, it should be mentioned again that outside a neighbourhood of $\tau = 0$ the proper time scale $t = f(\tau)$ is appropriate, at least for practical purposes.

4.3 The scale factor

At a first glance, the property $S(\tau) < 0$ for $\tau < 0$ could look strange. But one should be aware that even in FRW the supposition $K(t) > 0$ for $t > 0$ is nothing but a comfortable convention. One could also work with the convention $K(t) < 0$ for $t > 0$. Also in EFRW there does not occur anything special. The reason is that many physical quantities, e.g., curvature and metric, depend only on S^2 or $|S|$. Other quantities, e.g., the Hubble parameter, have to be redefined properly.

4.4 The horizon problem

The results of Section 3 formulated in Section 3.2.4 can also be expressed in the following form.

A model of EFRW does not have horizons exactly if it is either periodic or aperiodic with $\Lambda_0 = 0$. Then, in the latter case necessarily $k = 0, -1$. Moreover, if $k = 0, -1$ periodicity occurs if and only if $\Lambda_0 < 0$.

This result suggests complementing EFRW with the additional condition

AC: *The solutions S, ψ, ω are either periodic or aperiodic with $\Lambda_0 = 0$.*

In order to discuss the consequences of the above result with respect to the horizon problem let us first give a short characterization of it.

If a relativistic theory, denoted Ψ , has symmetries, e.g., homogeneity and isotropy, two questions arise.

1. Are there horizons in Ψ ?
2. If there are horizons in Ψ , why is it that the manifold M of Ψ exhibits symmetries, i.e., why have causally separated regions of M the same or very similar properties?

These two questions constitute what is called horizon problem in the present context. Since in EFRW plus AC the answer to the first question is “No”, so that the second question is empty, the horizon problem has been solved.

But one should be aware that there is a problem behind the horizon problem, which has not vanished with vanishing horizons. For, the nonexistence of horizons only guarantees causal contact for each particle with all other

particles in the world. Intuitively, this is a necessary condition for the large scale homogeneity and isotropy. But is it also sufficient? Hence, the horizon problem has a successor, the problem of large scale stability of universes with full causal contact of all particles. As far as I can see this problem is unsolved. However, calling it also “horizon problem” is abuse of language. This is because the term “horizon problem” has become an attribute of elements of the null set.

4.5 Outlook

Looking back at EFRW as it is presented in Section 2.5 one may notice that the theory contains unspecified elements, the constitutive quantities p and e , and the cosmological constant Λ_0 . Only occasionally some specifications were introduced. It was assumed in Section 3 that the universe near big crunch bang is asymptotically radiation dominated and that it is asymptotically matter dominated and $p = 0$ (i.e., matter is dust) for sufficiently large $|\tau|$. Moreover, in Section 4.4 the condition AC was introduced, which amounts roughly to $\Lambda_0 \leq 0$. This was done on purpose! It was not intended to study specified cosmological scenarios, rather to formulate a frame theory for possible scenarios.

In this way, EFRW is open to various combinations, especially all such scenarios which are used in FRW, such as inflation, variable speed of light etc. (cf. e.g., [10,17,18]) and the literature quoted there.)

But EFRW plus AC has the advantage that it avoids some severe shortcomings of FRW as it was demonstrated in this paper. This is a strong hint to cast out all models with open space-times and positive cosmological constants.

Acknowledgments

I want to thank my colleagues Mr. Gerhard Lessner for valuable discussions and critical remarks and Mr. Rolf Breuer for correcting my English.

References

- [1] A. Friedmann, *Über die Krümmung des Raumes*, Z. Phys. **10** (1922), 377.

- [2] A. Friedmann, *Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes*, Z. Phys. **21** (1924), 326.
- [3] H.P. Robertson, *On the Foundation of Relativistic Cosmology*, Proc. Natl. Acad. Sci. USA **15** (1929), 822.
- [4] H.P. Robertson, *Relativistic Cosmology*, Rev. Mod. Phys. **5** (1933), 62.
- [5] A.G. Walker, *Completely symmetric spaces*, J. Lond. Math. Soc. **19** (1944), 219.
- [6] St. Weinberg, *Gravitation and cosmology*, Wiley, New York, 1972.
- [7] W. Rindler, *Essential relativity*, Springer-Verlag, Heidelberg, Berlin, New York, 1977.
- [8] W. Rindler, *Visual Horizons in World-Models*, Mon. Not. R. Astron. Soc. **116** (1956), 662.
- [9] G. Boerner, *The early universe*, Springer-Verlag, Heidelberg, Berlin, New York, 1988.
- [10] E. Kolb and M.S. Turner, *The early universe*, Addison-Wesley Publ., 1990.
- [11] S.W. Hawking and G.F.R. Ellis, *The large scale structure of space-time*, Cambridge University Press, Cambridge, New York, 1973.
- [12] H. Goenner, *Einführung in die Kosmologie*, Spektrum Akad. Verlag, Heidelberg, Berlin, Oxford, 1994.
- [13] G. Lessner, *An open ever-oscillating universe*, Nuovo Cimento B **116** (2001), 59.
- [14] V. Perlick, *Characterization of Standard Clocks by Means of Light Rays and Freely Falling Particles*, Gen. Rel. Gravit. **19** (1986), 1059.
- [15] W. Kundt and B. Hoffmann, *Recent developments in general relativity*, Pergamon, Polish Scientific Publ., 1962.
- [16] R.F. Marzke and J.A. Wheeler, *Gravitation and relativity*, eds. H.Y. Chin and W.F. Hoffmann, W.A. Benjamin, New York, 1964.
- [17] A. Albrecht and J. Magueijo, *Time varying speed of light as a solution to cosmological puzzles*, Phys. Rev. D **59** (1999), 043516.
- [18] M. Gasperini and G. Veneziano, *The pre-big bang scenario in string cosmologies*, Phys. Rep. **373** (2003), 1.
- [19] R. Penrose, *General relativity*, eds. S.W. Hawking and W. Israel, Cambridge University Press, Cambridge, New York, 1979.
- [20] R.P. Geroch, *Relativity*, eds. S. Fickler, M. Carmeli and L. Witten, Plenum Press, New York, London, 1970.
- [21] F.J. Tipler, C.J.S. Clarke and G.F.R. Ellis, *General relativity and gravitation*, ed. A. Held, Plenum Press, New York, London, 1980.

