

On the genericity of spacetime singularities

PANKAJ S JOSHI

Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India
E-mail: psj@tifr.res.in

Abstract. We consider here the genericity aspects of spacetime singularities that occur in cosmology and in gravitational collapse. The singularity theorems (that predict the occurrence of singularities in general relativity) allow the singularities of gravitational collapse to be either visible to external observers or covered by an event horizon of gravity. It is shown that the visible singularities that develop as final states of spherical collapse are generic. Some consequences of this fact are discussed.

Keywords. Spacetime singularities; gravitational collapse; cosmology.

PACS Nos 04.20.Dw; 04.70.-s; 04.70.Bw

1. Introduction

After the advent of singularity theorems due to Hawking, Penrose and Geroch in the late nineteen sixties, the role of spacetime singularities as an inevitable feature of Einstein's theory of gravitation became clear. Such singularities occur in cosmology and in gravitational collapse. They also lead to situations where the gravitational field becomes ultra-strong and grows without any upper bound. The above-mentioned theorems further prove that the singularities manifest themselves in terms of the incompleteness of non-space-like geodesics in spacetime.

It is possible that such singularities represent the incompleteness of the theory of general relativity itself. Further, they may be resolved or avoided when quantum effects near them are included in a more complete theory of quantum gravity. Nevertheless, there is a key point here. Even if the final singularity is dissolved by quantum gravity, what is important is the inevitable occurrence of an ultra-strong gravity regime, close to the location of the classical singularity, in cosmology as well as during the dynamical processes involved in gravitational collapse. This affects the physics of the Universe. An example of such a situation is the big bang singularity of cosmology. Even though such singularities may be resolved through either quantum gravity effects [1], or features such as chaotic initial conditions, the effects of the super ultra-dense region of gravity that existed in the big bang epoch profoundly influence the physics and subsequent evolution of the Universe. Similarly, in the gravitational collapse of a massive star, the same theorems predict the occurrence of singularities which can be either visible to external observers or

hidden behind the event horizon giving rise to a black hole. In the case of a collapse process leading to a visible or naked singularity, again the important issue is how the inevitably occurring super ultra-dense region would influence the physics outside.

The above discussion does not imply the absence of singularity-free solutions to Einstein's equations. There are many examples of such solutions (see e.g. [2] in this volume). In fact, the Gödel Universe is a cosmological model which is geodesically complete. Also, Minkowski spacetime is the vacuum solution which is geodesically complete. For an additional discussion of singularity-free models in general relativity, we refer to [3]. Again, in gravitational collapse, one could start with a matter cloud which is initially collapsing. However, there could be a bounce later and the cloud may then disperse away, at least in principle. In such a case, no spacetime singularity or ultra-strong gravity region needs to form.

Since its derivation in the early 1950s, the Raychaudhuri equation [4] has played a central role in the analysis of spacetime singularities in general relativity. Prior to the use of this equation to analyse collapsing and cosmological situations for the occurrence of singularities [5], most works on related issues had taken up only rather special cases with many symmetry conditions assumed on the underlying spacetime. Raychaudhuri, however, considered for the first time these aspects within the framework of a general spacetime without any symmetry conditions, in terms of the overall behaviour of the congruences of trajectories of material particles and photons propagating and evolving dynamically. This analysis of the congruences of non-space-like curves, either geodesic or otherwise, showed how gravitational focusing took place in the Universe, giving rise to caustics and conjugate points. Before general singularity theorems could be constructed, however, another important mathematical input was needed in addition to the Raychaudhuri equation. This was the analysis of the causality structure and general global properties of a spacetime manifold. This particular development took place mainly in the late 1960s (for a detailed discussion, see e.g. [6] or [7]). The work of Penrose, Hawking, and Geroch then combined these two important features, namely the gravitational focusing effects due to matter and causal structure constraints following from global spacetime properties, to obtain the singularity theorems.

We review these developments briefly in §2. We then point out that the main question here is that of genericity of such spacetime singularities, either in cosmology or in collapse situations. The singularity theorems, while proving the existence of geodesic incompleteness, by themselves provide no information either on the structure and properties of such singularities or on the growth of curvature in their vicinity. Some of these issues, together with the possible avoidance of singularities, are considered in §3. We then discuss the genericity aspects of visible singularities in §4. Section 5 contains some concluding remarks.

2. The occurrence of singularities

We discuss here the occurrence of spacetime singularities in some detail within a general spacetime framework, and point out the crucial role of the Raychaudhuri equation therein. The basic ideas involved in the proofs of the singularity theorems are reviewed, and what these theorems do not imply is pointed out.

We observe the Universe today to very far depths in space and time. While looking deep into space, say for example in diametrically opposite directions, regions with extremely distant galaxies are seen in each direction. Interestingly, these regions have quite similar properties in terms of their appearance and the homogeneity in their galactic spatial distribution. These regions are, however, so far away from each other that they have had no time to interact mutually. That is because the age of the Universe since the big bang has not been large enough for such interactions to have occurred in the past. Thus, within the big bang framework of cosmology, a relevant question arises: how come these regions have such similar properties? This is one of the major puzzles of modern cosmology today.

This observed homogeneity and isotropy of the Universe at large enough scales can be modelled by the Robertson–Walker geometry. The metric of the corresponding spacetime is given by

$$ds^2 = -dt^2 + R^2(t) \left[\frac{dr^2}{(1 - kr^2)} + r^2 d\Omega^2 \right]. \quad (1)$$

Here $d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2$ is the metric on a two-sphere. The Universe is assumed to be spherically symmetric here. There is an additional assumption that the matter content of the spacetime is homogeneous and isotropic. That is, the matter density is the same everywhere in the Universe and also looks the same in all directions. Combining this geometry with the Einstein equations and solving the same, one is led to the Friedmann solution yielding a description of the dynamical evolution of the Universe. The latter requires that the Universe must have had a beginning at a finite time in the past. This is the epoch of the so-called big bang singularity. The matter density as well as the curvatures of spacetime diverge in the limit of approaching this cosmological singularity. This is where all non-spacelike geodesics are incomplete at a point in the past where spacetime comes to an end.

A similar occurrence of the formation of a spacetime singularity takes place when a massive star collapses freely under the force of its own gravity after the exhaustion of its nuclear fuel. If the mass of the star is small enough, the latter can stabilize as a white dwarf or a neutron star as its end-state. However, in case the mass is much larger, say of the order of tens of solar masses, a continual collapse is inevitable once there are no internal pressures left to sustain the star. Such a scenario was considered and modelled by Oppenheimer and Snyder [8], when they considered a collapsing spherical cloud of dust. Again, according to the equations of general relativity, a spacetime singularity of infinite density and curvature forms at the center of the collapsing cloud.

These singularities, after their discovery, were debated extensively by gravitation theorists. An important question that was persistently asked at this juncture was the following. Why should these models be taken so seriously when they assume so many symmetries? Perhaps such a spacetime singularity was arising as a result of these assumptions and could only occur in such a special situation. In other words, these singularities could be just some isolated examples. After all, the Einstein equations

$$R_{ab} - \frac{1}{2} R g_{ab} = 8\pi T_{ab}, \quad (2)$$

governing the ever present force of gravity, are a complex system of second-order, non-linear, partial differential equations admitting an infinite space of solutions in which the models discussed above are isolated examples. In other words, the issue was the absence of any proof that singularities would always occur in a general enough gravitational collapse when a massive star dies, or in a generic enough cosmological model. In fact, there was widespread belief in the 1940s and 1950s that such singularities would be removed both from stellar collapse and from the cosmological beginning of the Universe (which are two very important physical situations, where gravity becomes a dominant force), once assumptions like the dust form of matter and the spherical symmetry of the model were relaxed and more general solutions to the Einstein equations were considered. This is where the work by Raychaudhuri on the gravitational focusing of matter or light in a spacetime, and that of Penrose, Hawking, and Geroch on the causal structure and global properties of a general spacetime, the culmination of which were the singularity theorems, became relevant. These theorems showed that spacetime singularities, such as those depicted in the above examples, manifested themselves in a large class of models under quite general physical conditions.

We now outline the basic idea and the chain of logic behind the proof of a typical singularity theorem, highlighting the role of the Raychaudhuri equation therein. There are several singularity theorems available which establish the non-space-like geodesic incompleteness for a spacetime under different sets of physical conditions. Each of these may be more relevant to one or the other specific physical situation, and may be applicable to different physical systems such as stellar collapse or the Universe as a whole. However, the most general of these is the Hawking–Penrose theorem [9], which is applicable to both the collapse situation and the cosmological scenario. Let us first briefly describe the main steps in the proof of this theorem. Using causal structure analysis, it is first shown that between certain pairs of events in spacetime there must exist time-like geodesics of maximal length. However, from both causal structure analysis and the global properties of a spacetime manifold M (which is assumed to satisfy the generic condition as well as a specific energy condition), it follows that a causal geodesic, which is complete with regard to both the future and the past, must contain pairs of conjugate points where the nearby null or time-like geodesics intersect. One is then led to a contradiction because the maximal geodesics mentioned above cannot contain any conjugate points, the existence of which would be against their maximality. Thus M itself must have non-space-like geodesic incompleteness. We restate the theorem more carefully and accurately below.

A spacetime (M, g) cannot be time-like and null geodesically complete if the following are satisfied:

- (1) $R_{ij}K^iK^j \geq 0$ for all non-space-like vectors K^i ;
- (2) the generic condition is satisfied, that is, every non-space-like geodesic contains a point at which $K_{[i}R_{j]el[m}K_n]K^eK^l \neq 0$, where K is the tangent to the non-space-like geodesic;
- (3) the chronology condition holds, and
- (4) there exists in M either a compact achronal set without edge, or a closed trapped surface, or a point p such that for all past directed null geodesics from p , eventually the expansion parameter θ must be negative.

The first condition above is an energy condition. All classical fields have been observed to satisfy a suitable positivity of energy requirement. The second condition is a statement that all non-space-like trajectories do encounter some non-zero matter or stress-energy density somewhere during their entire path. The third is a global causality requirement to the effect that there are no closed time-like curves in the spacetime. Finally, in the last condition above, the first or the second part corresponds to a gravitational collapse situation, whereas the last part corresponds to gravitational focusing within a cosmological framework, where θ denotes the expansion for the congruence of non-space-like curves. If any of these conditions are satisfied then the theorem can go through.

The main idea of the proof is that one shows that the following three conditions cannot hold simultaneously: (a) every inextendible non-space-like geodesic contains pairs of conjugate points, (b) the chronology condition holds, (c) there exists an achronal set \mathcal{S} in spacetime such that $E^+(\mathcal{S})$ or $E^-(\mathcal{S})$ is compact.

In the above, for any set \mathcal{S} , the future of the set is denoted by $I^+(\mathcal{S})$, which is the set of all those points which can be connected by future-directed time-like curves from any point in \mathcal{S} . The past of a set, denoted by I^- is defined similarly dually. An achronal set is a set of spacetime events, of which no two are chronologically related. The set $E(\mathcal{S})$ here denotes the edge of \mathcal{S} which is a set of all points $x \in \mathcal{S}$ such that every neighbourhood of x contains $y \in I^+(x)$ and $z \in I^-(x)$, such that there is a time-like curve that does not meet \mathcal{S} .

If the above is shown then the theorem is proved, because the condition (3) is same as (b), the condition (4) implies (c), and conditions (1) and (2) imply (a). First, we note that (a) and (b) imply the strong causality of the spacetime (see Proposition 6.4.6 of [5]). Strong causality is a causality condition stronger than the chronology which rules out the occurrence of closed time-like curves, and ensures that once a time-like path has left an event in the spacetime, it cannot return in the arbitrary vicinity of the same.

While for further details and definitions, we refer to [5] or [6], in the following we give a brief outline of the main argument for the proof. It can be shown that if \mathcal{S} is a future trapped set, which essentially means that no non-space-like paths escape its own future, and if strong causality holds on $\bar{I}^+(\mathcal{S})$ then there exists a future endless trip γ such that $\gamma \subset \text{Int } D^+(E^+(\gamma))$. Here the set D^+ denotes the future domain of dependence for a set. Now, one defines $T = \bar{J}^-(\gamma) \cap E^+(\mathcal{S})$, then T turns out to be past trapped and hence there exists λ , a past endless causal geodesic in $\text{Int}(D^-(E^-(T)))$. Then one chooses a sequence $\{a_i\}$ receding into the past on λ and a sequence $\{c_i\}$ on γ to the future. The sets $J^-(c_i) \cap J^+(a_i)$ are compact and globally hyperbolic. So there exists a maximal geodesic μ_i from a_i to c_i for each i . The intersections of μ_i with the compact set T have a limit point p and a limiting causal direction. The causal geodesic μ with this direction at p must have a pair of conjugate points. This is then shown to be contradictory to the maximality property of the geodesics stated above.

The Raychaudhuri equation plays a key role in the above argument. It implies the occurrence of conjugate points on the null or time-like geodesics under consideration. To state this in a somewhat different way, let us suppose that the spacetime manifold M is non-space-like geodesically complete. In that case, causal structure and the causality as well as regularity assumptions imply that there must exist

maximal non-space-like geodesics between certain pairs of spacetime points or along the non-space-like geodesics which are orthogonal either to a space-like hypersurface or to a trapped spacetime surface. However, a suitable energy condition, such as the weak energy condition given by $T_{ab}V^aV^b \geq 0$, together with Einstein equations, implies that $R_{ab}V^aV^b \geq 0$, which in turn implies from the Raychaudhuri equation that conjugate points must develop along these maximal length geodesic trajectories. This contradicts the statement, dictated by causal structure requirements, that maximal geodesics cannot contain any conjugate points.

3. Singularity avoidance

It would be quite attractive if one could avoid any such singularity at the classical level itself. Efforts were made in that direction after singularities had been found in the special cases of FRW cosmologies and the Schwarzschild solution related to gravitational collapse. It was believed that, by going to situations which did not require those exact symmetries, one might get rid of spacetime singularities altogether. However, as discussed above, the singularity theorems show that this is not possible. Once we accept that geodesic incompleteness represents a genuinely singular behaviour, the only way to avoid spacetime singularities at a purely classical level within the framework of Einstein gravity would be to somehow break or violate one of the assumptions or conditions which have been used in proving these theorems. Let us consider this possibility in some detail. The assumptions used in proving various singularity theorems fall mainly into three classes, as pointed out above. The first is a reasonable causality condition on spacetime; it ensures an overall well-behaved global behaviour for the Universe. The second is some suitable energy condition that emphasises the positive nature of the energy density of classical matter fields. The final assumption is the occurrence of a suitable focusing effect in spacetime, as caused by the presence of matter fields. This occurrence is typically in the form of trapped surfaces present in situations of gravitational collapse or a suitable overall focusing in the cosmological situation.

First, consider a possible violation of the causality condition. Note that the Einstein equations as such do not demand causality of spacetime as a whole. In fact, they impose no constraints on the global topology of spacetime. Hence, without breaking any consistency requirements, one could consider causality violating spacetimes and see if somehow the occurrence of spacetime singularities could be avoided there. This issue has been analysed in some detail. It turns out that if causality is violated in any finite region of spacetime, that in its own right causes spacetime singularities in the form of geodesic incompleteness [10,11]. These results show that the occurrence of closed time-like lines in any finite region of spacetime necessarily causes singularities there. One could of course preserve causality in the Universe by disallowing closed non-space-like curves, but still admitting higher-order causality violations. The latter can occur from time-like paths coming back to arbitrarily close neighbourhoods of an event that they had left. Or, a spacetime (M, g) could be causal, but allowed to admit closed time-like curves with the slightest perturbation of the metric. This last situation is described as a violation of the stable causality condition. The above results show that such higher-order causality

violations also give rise to spacetime singularities. One may then conclude that violating causality is not an effective alternative means to avoid spacetime singularities. Of course, as noted earlier, the Gödel Universe is geodesically complete and violates causality at every event of spacetime. So there might be a possibility that the violation of causality, introduced at every point of spacetime, may be able to avoid any singularity. However, such a Universe itself would be quite ‘singular’ in some sense. Indeed, because of this and other unphysical features, the Gödel Universe has never been regarded as a suitable model [5].

The second alternative, of violating the energy condition, has again been investigated in some detail (see e.g. [10] for a discussion on this). The essential conclusion here has been the following. So long as this condition, even if violated locally, holds in a spacetime averaged sense, the existence of singularities in the form of non-space-like geodesic incompleteness would be implied. Only if one violates the energy condition for almost all the events in the Universe making the energy density an overall negative quantity, could one avoid any possible singularity. This does not look to be physically achievable, at least for classical fields. So one could conclude that the violation of the energy condition does not offer an effective alternative in avoiding spacetime singularities. In other words, local violations of the energy condition do not affect the occurrence of singularities. Furthermore, no physical mechanisms have been observed or known which may allow such a violation of the energy condition in the Universe.

This brings us to the final and most important alternative, namely that of avoiding sufficient gravitational focusing in spacetime by breaking condition (4) above in the proof of the singularity theorem. This would amount to disallowing the formation of any trapped surfaces in gravitational collapse and similarly not having enough convergence in a cosmological scenario. Basically, if one wants to avoid all trappings, without violating the energy condition, the only option would be to have very little matter and stress–energy density, so that light rays just avoid getting sufficiently focused. This is the direction that some recent works appear to point to [2,12]. Similar results were obtained earlier, showing that if the space-like surfaces have sufficient matter in some non-vanishing average sense, then *all* non-space-like trajectories would be incomplete in the past [13,14]. For example, if the microwave background radiation is taken to be universal and also to have a global minimum in its energy density over a space-like surface, that will cause the necessary convergence; all non-space-like trajectories will then be incomplete in the past. Thus the avoidance of trapped surfaces or cosmological focusing would be one way not to have space-time singularities. However, if trapped surfaces did form in spherical or other more general gravitational collapse processes, the occurrence of singularities will be a generic phenomenon. We can elaborate this last point. Let such trapped surfaces develop as the collapsing system evolves dynamically. Could we generalize these conclusions for a non-spherically symmetric collapse? Are they valid at least for small perturbations from exact spherical symmetry. By using [5] the stability of Cauchy development in general relativity, one can show that the formation of trapped surfaces is a stable feature when departures from spherical symmetry are taken into account. The argument goes as follows. Consider a spherically symmetric collapse evolution from given initial data on a partial Cauchy surface S . Then trapped surfaces \mathcal{T} are found to form in the shape of all spheres with $r < 2m$ in the

exterior Schwarzschild geometry. The stability of Cauchy development then implies that, for all initial data sufficiently near the original data in the compact region $J^+(S) \cap J^-(\mathcal{T})$, trapped surfaces must still occur. The curvature singularity of a spherical collapse also turns out to be a stable feature as implied by the singularity theorems discussed above. The latter shows that closed trapped surfaces always imply the existence of a spacetime singularity under reasonably general conditions. In this sense, such singularities, when they occur, are really generic.

4. Genericity of naked singularities in spherical collapse

The above consideration points to a wide variety of circumstances under which singularities develop in general relativistic cosmologies and in many gravitational collapse processes. Singularity theorems imply the existence of vast classes of solutions to the Einstein equations that must contain spacetime singularities, as characterized by the conditions of these theorems, and of which the big bang singularity is one example. These theorems therefore imply that singularities must occur in Einstein's theory quite generically, i.e. under rather general physically reasonable conditions on the underlying spacetime. Historically, this implication considerably strengthened our confidence in the big bang model which is used extensively in cosmology today.

As mentioned earlier, the singularities are predicted to be either visible to external observers or hidden inside the event horizons of gravity. This is related to the causal structure in the vicinity of the concerned singularity and is particularly relevant to realistic gravitational collapse scenarios and to the basics of black hole physics. However, the singularity theorems provide no further information. Yet, we need more information on the structure of the singularities in terms of their visibility, curvature strengths and other such aspects. What is therefore called for is a detailed investigation of the dynamics of gravitational collapse within the framework of Einstein's theory. Extensive investigations have been made in the past couple of decades or so from such a perspective. These have dealt with dynamical gravitational collapse, mainly for spherical models but also in some non-spherical cases, investigating the structure and visibility of the concerned singularities. Depending on the initial conditions and the types of evolution of the collapse process allowed by the Einstein equations, the singularities of collapse can be either visible or covered (see, for example, [15–20] for some recent reviews and further references). For a detailed discussion of the gravitational collapse of radiation shells within a Vaidya metric and the related black hole and naked singularity geometries, we refer to [7].

Based on the work on spherical gravitational collapse done in [21–23], we consider here the genericity of naked singularities forming in such a process. Given the data on the initial density and pressure profiles of the collapsing cloud, classes of solutions to Einstein equations have been constructed which evolve to either a visible or a covered singularity, subject to the satisfaction of regularity and energy conditions. In other words, given the initial matter data on a space-like surface from which the collapse of the massive matter cloud evolves, the rest of the free functions, such as the velocities of the collapsing shells as well as the classes of evolution,

can be chosen as allowed by the Einstein equations, subject to an energy condition and suitable regularity conditions. The latter take the collapse either to a black hole or a naked singularity in the final state, depending on the choice made. The basic formalism here can be summarized as follows. The form of matter considered is quite generic, which is any Type-I general matter field subject to an energy condition [5]. In the case of a black hole developing as the end-state of collapse, the spacetime singularity is necessarily hidden behind the event horizon of gravity. In contrast, a naked singularity develops when families of future directed non-space-like trajectories come out from the singularity. These trajectories can in principle communicate information to faraway observers in the spacetime. The existence of such families confirms the formation of a naked singularity, as opposed to a black hole end-state.

The eventual singularity, which is the singularity curve produced by the collapsing matter, is what we study. We show here that the tangent to this curve at the central singularity at $r = 0$ is related to the radially outgoing null geodesics from the singularity, if any. By determining the nature of the singularity curve and its relation to the initial data and the classes of collapse evolution, we are able to deduce whether the formation of any trapped surface during the collapse takes place before or after the singularity. It is this causal structure of the trapped region that determines the possible emergence or otherwise of non-space-like paths from the singularity. That then settles the final outcome of the collapse in terms of either a black hole or naked singularity. Several familiar equations of state for which extensive collapse studies have been made, such as dust or matter with only tangential or radial pressures and others, are special cases of our consideration.

The spacetime geometry within the spherically symmetric collapsing cloud is described by the general metric in the comoving coordinates (t, r, θ, ϕ) as

$$ds^2 = -e^{2\nu} dt^2 + e^{2\psi} dr^2 + R^2(t, r) d\Omega^2, \quad (3)$$

where $d\Omega^2$ is the line element on a two-sphere. The matter fields, considered by us, belong to a broad class, called Type-I, where the energy-momentum tensor has one time-like and three space-like eigenvectors. This general class includes most physically reasonable matter, including dust, perfect fluids, massless scalar fields and such others. The stress-energy tensor for this class is given in a diagonal form: $T_t^t = -\rho, T_r^r = p_r, T_\theta^\theta = T_\phi^\phi = p_\theta$. Here ρ , p_r and p_θ are the energy density, the radial pressure and the tangential pressure respectively. We also take the matter field to satisfy the weak energy condition. This means that the energy density measured by any local observer must be non-negative. So, for any time-like vector V^i , one must have $T_{ik} V^i V^k \geq 0$. The latter amounts to taking $\rho \geq 0$, $\rho + p_r \geq 0$, $\rho + p_\theta \geq 0$. Now, for the metric (3), Einstein equations take the form (in the units $8\pi G = c = 1$):

$$\rho = \frac{F'}{R^2 \dot{R}}; \quad p_r = -\frac{\dot{F}}{R^2 \dot{R}}, \quad (4)$$

$$\nu' = \frac{2(p_\theta - p_r)}{\rho + p_r} \frac{R'}{R} - \frac{p_r'}{\rho + p_r}, \quad (5)$$

$$-2\dot{R}' + R' \frac{\dot{G}}{G} + \dot{R} \frac{H'}{H} = 0, \quad (6)$$

$$G - H = 1 - \frac{F}{R}, \quad (7)$$

where we have defined $G(t, r) = e^{-2\psi}(\dot{R}')^2$ and $H(t, r) = e^{-2\nu}(\dot{R})^2$.

The arbitrary function $F = F(t, r)$ has the interpretation of the mass function for the cloud, giving the total mass in a shell of comoving radius r . The energy condition implies that $F \geq 0$. In order to preserve regularity at the initial epoch, we need to take $F(t_i, 0) = 0$, i.e. the mass function must vanish at the center of the cloud. As seen from eq. (4), there is a density singularity in the spacetime at $R = 0$, and one at $R' = 0$. However, the latter is due to shell-crossings and can be possibly removed [24] through a suitable extension of the spacetime. In any case, we are interested here only in the shell-focusing singularity at $R = 0$. This is a physical singularity where all the matter shells collapse to a zero physical radius. We can use the scaling freedom available for the radial coordinate r to write $R = r$ at the initial epoch $t = t_i$ from where the collapse commences. Introducing the function $v(t, r)$ by the relation

$$R(t, r) = rv(t, r), \quad (8)$$

we have $v(t_i, r) = 1$ and $v(t_s(r), r) = 0$. The collapse situation is now characterized by the condition $\dot{v} < 0$. The time $t = t_s(r)$ corresponds to the shell-focusing singularity $R = 0$ where all the matter shells collapse to a vanishing physical radius.

From the standpoint of the dynamic evolution of the initial data, prescribed at the initial epoch $t = t_i$, there are six arbitrary functions of the comoving shell radius r as given by, $\nu(t_i, r) = \nu_0(r)$, $\psi(t_i, r) = \psi_0(r)$, $R(t_i, r) = r\rho(t_i, r) = \rho_0(r)$, $p_r(t_i, r) = p_{r_0}(r)$, $p_\theta(t_i, r) = p_{\theta_0}$. We note that not all of the initial data above are mutually independent, because from eq. (5) one gets a relation that gives the initial function $\nu_0(r)$ in terms of the rest of the other initial data functions. The initial pressures should have a physically reasonable behaviour at the center ($r = 0$) in that the pressure gradients should vanish there, i.e. $p'_{r_0}(0) = p'_{\theta_0}(0) = 0$, and also the difference between the radial and the tangential pressures needs to vanish at the center, i.e. $p_{r_0}(0) - p_{\theta_0}(0) = 0$. These conditions are necessary to ensure the regularity of the initial data at the center of the cloud. It is then evident from eq. (9) that $\nu_0(r)$ has the form

$$\nu_0(r) = r^2 g(r), \quad (9)$$

where $g(r)$ is at least a C^1 function of r for $r = 0$, and at least a C^2 function for $r > 0$. We see that there are five total field equations with seven unknowns, ρ , p_r , p_θ , ψ , ν , R , and F . Thus we have the freedom to choose two free functions. In general, their selection, subject to the weak energy condition and the given initial data for collapse at the starting surface, determines the matter distribution and the metric of the spacetime, and thus leads to a particular time evolution of the initial data.

Let us construct the classes of solutions to Einstein equations which give the collapse evolution from the given initial data. Consider first the general class of mass functions $F(t, r)$ for the collapsing cloud. These are given as

$$F(t, r) = r^3 \mathcal{M}(r, v), \quad (10)$$

\mathcal{M} being any regular and suitably differentiable general function without further restrictions. It turns out from equations given below that the regularity and finiteness of the density profile at the initial epoch $t = t_i$ requires F to go as r^3 close to the center. It follows that the form of F above is not really any special choice, but is in the general class of mass functions consistent with the collapse regularity conditions. Equations (4) yield

$$\rho = \frac{3\mathcal{M} + r[\mathcal{M}_{,r} + \mathcal{M}_{,v}v']}{v^2(v + rv')}; \quad p_r = -\frac{\mathcal{M}_{,v}}{v^2}. \quad (11)$$

Thus the regular density distribution at the initial epoch is given by $\rho_0(r) = 3\mathcal{M}(r, 1) + r\mathcal{M}_{,r}(r, 1)$. It is evident in general that, as $v \rightarrow 0$, $\rho \rightarrow \infty$ and $p_r \rightarrow \infty$, i.e. both the density and the radial pressure blow up at the singularity. One can, in fact, show the following. Given any regular initial density and pressure profiles for the matter cloud from which the collapse develops, there always exist velocity profiles for collapsing matter shells as well as classes of dynamical evolutions, as determined by the Einstein equations, which lead to a naked singularity or a black hole as the end-state of collapse, depending on the choice of the class.

We now provide classes of solutions to Einstein equations to this effect. Consider the class of velocity profiles as determined by the general function

$$\nu = A(t, R), \quad (12)$$

where $A(t, R)$ is any arbitrary, suitably differentiable function of t and R , with the initial constraint $A(t_i, R) = \nu_0(r)$. Using (12) in (6), we have

$$G(t, r) = b(r)e^{2(A - \int A_{,t} dt)}. \quad (13)$$

Here $b(r)$ is another arbitrary function of r which emerges on integrating the Einstein equations. A comparison with dust models enables one to interpret $b(r)$ as the velocity function for the shells. From eq. (9), the form of $A(t, R)$ is seen to be $A(t, R) = r^2 g_1(r, v)$, where $g_1(r, v)$ is a suitably differentiable function and $g_1(r, 1) = g(r)$. Similarly, we have $A - \int A_{,t} dt = r^2 g_2(r, v)$ and at the initial epoch $g_2(r, 1) = g(r)$. Using (12) in (5), we obtain

$$2p_\theta = RA_{,R}(\rho + p_r) + 2p_r + \frac{Rp'_r}{R}. \quad (14)$$

In general, both the density and the radial pressure are known to blow up at the singularity. However, the above equation implies that the tangential pressure also does the same.

Writing

$$b(r) = 1 + r^2 b_0(r) \quad (15)$$

and using eqs (10), (12) and (13) in (7), we are led to

$$\sqrt{R}\dot{R} = -e^{r^2 g_1(r,v)} \sqrt{(1 + r^2 b_0) R e^{r^2 g_2(r,v)} - R + r^3 \mathcal{M}}. \quad (16)$$

Since we are considering a collapse situation, we have $\dot{R} < 0$. Defining a function $h(r, v)$ as

$$h(r, v) = \frac{e^{2r^2 g_2(r,v)} - 1}{r^2} = 2g_2(r, v) + \mathcal{O}(r^2 v^2) \quad (17)$$

and using (17) in (16), we obtain after some simplifications:

$$\sqrt{v}\dot{v} = -\sqrt{e^{2r^2(g_1+g_2)} v b_0 + e^{2r^2 g_1} (v h(r, v) + \mathcal{M}(r, v))}. \quad (18)$$

Integrating the above equation we have

$$t(v, r) = \int_v^1 \frac{\sqrt{v} dv}{\sqrt{e^{2r^2(g_1+g_2)} v b_0 + e^{2r^2 g_1} (v h + \mathcal{M})}}. \quad (19)$$

Note that r is treated as a constant in the above integration. Expanding $t(v, r)$ around the center, we have

$$t(v, r) = t(v, 0) + r\mathcal{X}(v) + \mathcal{O}(r^2), \quad (20)$$

where the function $\mathcal{X}(v)$ is given by

$$\mathcal{X}(v) = -\frac{1}{2} \int_v^1 dv \frac{\sqrt{v}(b_1 v + v h_1(v) + \mathcal{M}_1(v))}{(b_{00} v + v h_0(v) + \mathcal{M}_0(v))^{3/2}}, \quad (21)$$

with, $b_{00} = b_0(0)$, $\mathcal{M}_0(v) = \mathcal{M}(0, v)$, $h_0 = h(0, v)$, $b_1 = b'_0(0)$, $\mathcal{M}_1(v) = \mathcal{M}_{,r}(0, v)$ and $h_1 = h_{,r}(0, v)$.

From the above, the time when the central singularity develops is given by

$$t_{s_0} = \int_0^1 \frac{\sqrt{v} dv}{\sqrt{b_{00} v + v h_0(v) + \mathcal{M}_0(v)}}. \quad (22)$$

The time for other shells to reach the singularity can be given by the expansion

$$t_s(r) = t_{s_0} + r\mathcal{X}(0) + \mathcal{O}(r^2). \quad (23)$$

It is now clear that the value of $\mathcal{X}(0)$ depends on the functions b_0 , \mathcal{M} and h , which in turn depend on the initial data at $t = t_i$ and on the dynamical variable v that evolves in time from a value $v = 1$ at the initial epoch to $v = 0$ at the singularity. Thus, a given set of initial matter distribution and the dynamical profiles including the velocity of shells completely determine the tangent at the center to the singularity curve. Equations (17)–(19), lead to

$$\sqrt{v}v' = \mathcal{X}(v) \sqrt{b_{00} v + v h_0(v) + \mathcal{M}_0(v)} + \mathcal{O}(r^2). \quad (24)$$

The end-state of collapse, in the form of either a black hole or a naked singularity, is determined by the causal behaviour of the apparent horizon, which is the boundary of trapped surfaces forming due to collapse, and is given by $R = F$. If the neighbourhood of the center gets trapped prior to the epoch of singularity, then it is covered and a black hole results. Otherwise, it would be a naked singularity with the non-space-like future directed trajectories escaping from it. It is then to be determined if there are any families of future directed non-space-like paths emerging from the singularity. To see this as well as to examine the nature of the central singularity at $R = 0$, $r = 0$, one has to consider the equation

$$\frac{dt}{dr} = e^{\psi-\nu} \quad (25)$$

for outgoing radial null geodesics. For a further discussion on this, we refer to [21–23]. The main result that follows from this is that if the quantity $\mathcal{X}(0)$ is positive, one indeed gets radially outgoing null geodesics coming out from the central singularity, which then has to be naked. However, if $\mathcal{X}(0)$ is negative, we have a black hole solution, since there are no such trajectories coming out. If $\mathcal{X}(0)$ vanishes, then we have to take into account the next higher-order non-zero term in the singularity curve $t_s(r)$, and a similar analysis can be carried out.

In the above discussion, the functions h and \mathcal{M} have been expanded in r around $r = 0$ and only the first-order terms have been retained. The key point here is that the functions chosen for the classes of collapse evolution from the given initial data are such that the eventual singularity curve $t_s(r)$ is expandable at the center $r = 0$, at least to first order. The well-known and extensively studied example of dust collapse (see e.g. [25] and references therein) does satisfy this feature. It thus forms a special sub-class of the above classes of solutions for spherical collapse. Various other collapse models studied earlier, such as collapse with a non-zero tangential pressure and others also satisfy such a behaviour of the singularity curve. For a detailed discussion of the differentiability conditions on the functions involved and of the classes of models that satisfy these, we refer to the papers cited above. The treatment given here applies to general classes of functions with the minimum differentiability conditions. However, in discussions of collapse these are sometimes assumed to be analytic and expandable with respect to r^2 with the argument that such smooth functions are physically more relevant. Such assumptions are in fact due to computational convenience, though one can mathematically exploit the freedom of definition and choice available. The formalism, outlined above, would of course work for such smooth functions, which constitute a special case of what has been considered. Consequently, one can have a smooth and differentiable singularity curve.

We thus see how the initial data determine the black hole and naked singularity phases as collapse end-states in terms of the available free functions. This happens because the quantity $\mathcal{X}(0)$ is determined by these initial and dynamical profiles, as given by (21). It is clear that, given any regular density and pressure profiles for the matter cloud from which the collapse develops, one can always choose velocity profiles such that the end-state of the collapse would be either a naked singularity or a black hole, and vice versa. It is interesting to note here that physical agencies, such as the spacetime shear within a dynamically collapsing cloud, could naturally give rise to such phases in gravitational collapse [26]. In other words, such physical

factors can naturally delay the formation of any apparent horizon and trapped surfaces.

As stated above, we have worked here with the Type-I matter fields. This is a rather general form of matter which includes practically all known physically reasonable fields such as dust, perfect fluids, massless scalar fields and so on. However, note that we have assumed no explicit equation of state relating the density and pressure variables. Assuming a specific equation of state will fully close the system, all the concerned functions being then determined as a result of the development of the system from the initial data. We presently have little idea about the kind of equation of state that the matter should follow. This is particularly true when the matter is close to the collapse end-state, where we are dealing with ultra-high densities and pressures. One might as well be allowed to freely choose the property of the matter fields, as done above. It is nevertheless important to make the following point. The analysis given above does include several well-known classes of collapse models and equations of state. We had a choice of two free functions available and constructed general classes of collapse evolution using the mass function $F(r, v)$ and the metric function ν . Suppose now that, in addition to Einstein equations, an equation of state of the form $p_r = f(\rho)$ (or alternatively, $p_\theta = g(\rho)$) is given. Then, as evident from eq. (4), there would be a constraint on the otherwise arbitrary function \mathcal{M} , specifying the required class, provided a solution to the constraint equation exists. Then the value of p_θ (or p_r) gets determined in terms of $\rho(t, r)$ from eq. (5), and the analysis for the black hole and naked singularity phases goes through as above.

For example, in the case of collapsing dust models, we have

$$p_r = p_\theta = 0 \quad (26)$$

and the constraint equation yields

$$\mathcal{M}(r, v) = \mathcal{M}(r). \quad (27)$$

Next, for collapse with a constant (or zero) radial pressure, but with an allowed variable tangential pressure, the constraint equation leads to

$$\mathcal{M}(r, v) = f(r) - kv^3, \quad (28)$$

where k is the value of the constant radial pressure. The tangential pressure will then be given by

$$p_\theta = k + \frac{1}{2}A_{,R}R(\rho + k). \quad (29)$$

Along with these two well-known models, the analysis works for any other model in which any one of the pressures is specified by an equation of state and which permits a solution to the constraint equation on \mathcal{M} . This provides some insight into how these black hole and naked singularity phases come about as collapse end-states.

The above results provide us with information on the genericity aspects of naked singularities developing as the final states of spherical collapse. Consider a specific

collapse evolution for which the quantity $\mathcal{X}(0)$ is positive, leading to the formation of a naked singularity as the end-state. As noted earlier, this value is decided by the initial density and pressure profiles and the class of dynamical evolutions as determined by the Einstein equations. The initial velocity profiles for the collapsing shells are controlled by the function $b_0(r)$ and the collapse evolutions $\mathcal{M}(r, v)$ and $h(r, v)$, which in turn depend on the initial data at $t = t_i$ and the dynamical variable v . Hence, by continuity, any arbitrarily small variation in any of these parameters will continue to yield a value for $\mathcal{X}(0)$ which is again positive. Thus if a collapse evolution creates a naked singularity, any sufficiently small variation in the initial matter and velocity profiles for the collapsing shell, or small variations in the dynamical evolutions functions will still preserve the final outcome of a naked singularity. In this sense the naked singularities of spherical collapse turn out to be completely generic.

We paraphrase the basic result discussed in this article. Given a starting set of regular data in terms of the density and pressure profiles at the initial epoch from which the collapse develops, there are sets of dynamical evolutions, and velocity profiles for the collapsing cloud, i.e. classes of solutions to the Einstein equations, which can evolve the given initial data to produce either a black hole or a naked singularity as the collapse end-state. The exact outcome depends on the choice of the rest of the free functions available. Indeed, the space of initial data can be divided into two distinct subspaces: one containing those that evolve into black holes and another containing those that lead to naked singularities. The latter are generic in the sense described above. In a way this brings out the stability of these collapse outcomes with respect to perturbations in the initial data set and in the choice of collapsing matter fields.

5. Concluding remarks

In this article we have discussed genericity aspects of spacetime singularities. It is seen that in Einstein gravity they occur generically, whether covered within event horizons or as visible to external observers. This is a consequence of the singularity theorems which make crucial use of the Raychaudhuri equation. It is possible that a future quantum theory of gravity may remove or resolve the final singularity of collapse or the initial one in cosmology. In such a scenario, what matters physically are the regions of ultra-strong gravity and spacetime curvature that develop as a result of the accompanying dynamical gravitational processes. This is true even if the very final singularity is removed through quantum effects. The following physical picture then emerges. Dynamical gravitational processes proceed and evolve to create ultra-strong gravity regions in the Universe. Once these form, strong curvature and quantum effects both come into their own in these regions. Quantum gravity then takes over and is likely to resolve the final singularity. Particularly interesting is the case when the singularities of collapse are visible. In such a situation, quantum gravity effects, taking place in those ultra-strong gravity regions, will in principle be accessible and observable to external observers. The consequences are likely to be intriguing [27].

Acknowledgment

I would like to express my deep gratitude to the late A K Raychaudhuri for his keen interest and extremely relevant comments on our work on gravitational collapse and spacetime singularities. It was a great pleasure knowing him and discussing with him, and to benefit from his penetrating insights. I like to warmly thank Probir Roy for several helpful comments, which clarified many important points.

References

- [1] See for example, M Bojowald (2005), Living Rev. Rel. 8:11; and references therein, for details on loop quantum gravity effects near the big bang singularity
- [2] J Senovilla, *Pramana – J. Phys.* **69**(1), 31 (2007), gr-qc/0610127
- [3] M Rayn and S Shepley, *Inhomogeneous cosmological models* (Princeton University Press, Princeton, 1975)
- [4] A K Raychaudhuri, *Phys. Rev.* **98**, 1123 (1955)
See also *Pramana – J. Phys.* **69**(1), 3 (2007). It gives an insight into how he arrived at his equation
- [5] S W Hawking and G F R Ellis, *The large scale structure of space-time* (Cambridge University Press, Cambridge, 1973)
- [6] R M Wald, *General relativity* (University of Chicago Press, Chicago, 1984)
- [7] P S Joshi, *Global aspects in gravitation and cosmology* (Clarendon Press, OUP, Oxford, 1993)
- [8] J R Oppenheimer and H Snyder, *Phys. Rev.* **56**, 455 (1939)
- [9] S W Hawking and R Penrose, *Proc. R. Soc. London* **A314**, 529 (1970)
- [10] F Tipler, C J S Clarke and G F R Ellis, in *General relativity and gravitation* edited by A Held (Plenum, New York, 1980) vol. 2
- [11] P S Joshi, *Phys. Lett.* **A85**, 319 (1981)
- [12] N Dadhich and A Raychaudhuri, *Mod. Phys. Lett.* **A14**, 2135 (1999)
- [13] P S Joshi, *Pramana – J. Phys.* **15**, 225 (1980)
- [14] P S Joshi, *Bull. Astron. Soc. India* **14**, 1 (1986)
- [15] A Krolak, *Prog. Theor. Phys. Suppl.* **136**, 45 (1999)
- [16] R Penrose in *Black holes and relativistic stars* edited by R M Wald (University of Chicago Press, 1998)
- [17] P S Joshi, *Pramana – J. Phys.* **55**, 529 (2000)
- [18] M Celerier and P Szekeres, gr-qc/0203094
- [19] T Harada, H Iguchi and K Nakao, *Prog. Theor. Phys.* **107**, 449 (2002)
- [20] R Giambo, F Giannoni, G Magli and P Piccione, *Commun. Math. Phys.* **235**, 545 (2003)
- [21] P S Joshi and I H Dwivedi, *Class. Quant. Grav.* **16**, 41 (1999)
- [22] P S Joshi and R Goswami, *Phys. Rev.* **D69**, 064027 (2004)
- [23] R Goswami and P S Joshi, gr-qc/0608136 (2006). A discussion on generalization of gravitational collapse scenarios to higher dimensional spacetimes, and the differences that are caused due to increase in dimensions is contained here
- [24] C J S Clarke, *The analysis of spacetime singularities* (Cambridge University Press, Cambridge, 1993)

- [25] P S Joshi and I H Dwivedi, *Phys. Rev.* **D47**, 5357 (1993)
- [26] P S Joshi, N Dadhich and R Maartens, *Phys. Rev.* **D65**, 101501 (2002). It is seen that while in the Vaidya collapse of radiation it is the time rate of collapse that governs the black hole and naked singularity final phases, in the dust case treated above it is the amount of inhomogeneity near the center that governs the final collapse outcome
- [27] R Goswami, P S Joshi and P Singh, *Phys. Rev. Lett.* **96**, 031302 (2006)