

Critical Behaviour in Gravitational Collapse

MICHAEL KINACH

Date: April 2019
Course: PHYS 530B
Instructor: Dr. W. Unruh

Abstract

Gravitational collapse exhibits some unexpected behaviour near the threshold of black hole formation: self-similarity, power-law scaling and universality. These three phenomena closely resemble the physics of second-order phase transitions in statistical mechanics and are therefore labelled *critical phenomena*. In this paper, I describe these phenomena in detail and provide some historical context for their discovery. I also explore the analogy between critical phenomena in gravitational collapse and in statistical mechanical systems.

1 Introduction

In general relativity, a smooth distribution of matter will typically evolve into one of three distinct end states:

1. It will disperse to infinity if the internal interactions are too weak to hold it together.
2. It will equilibrate to a stable star if an outward force (such as thermal pressure) balances the inward gravitational attraction.
3. It will collapse to a black hole if the gravitational attraction becomes too strong to overcome.

Given these possibilities, it is worthwhile to ask how a generic matter distribution will evolve into one of the three end states. It is well known that for a sufficiently weak initial configuration of matter, the system will disperse and evolve into Minkowski spacetime [1]. Likewise, a sufficiently strong configuration will collapse into a black hole [2]. However, there remains some ambiguity in the intermediate regime when the matter is teetering on the cusp of gravitational collapse. This raises an interesting question: what is the exact threshold of black hole formation?

This question was first proposed by Demetrios Christodoulou to Matthew Choptuik in the spring of 1987 [3]. Both were studying the spherically-symmetric collapse of a minimally coupled massless scalar field. Christodoulou was considering the scalar field model in terms of single-parameter families of spacetimes characterized by some parameter p (for example, p could be the amplitude of the scalar field). For a given family, the parameter p can be varied so that the scalar field will either disperse to infinity or collapse to a black hole. We define some critical value p^* as the threshold where black hole formation suddenly “turns on”. Christodoulou’s question was whether this critical black hole would have finite or infinitesimal mass.

Using numerical methods, Choptuik answered this question: black holes at the critical threshold can form with *infinitesimal* mass [4]. His analysis also revealed some unexpected behaviour. Notably, he found that the mass M of a black hole scales as

$$M \simeq (p - p^*)^\gamma \tag{1}$$

where $\gamma \approx 0.37$. This relation appears to hold for any choice of the parameter p and for many different families of initial data. Secondly, he found that the scalar field near the critical threshold exhibits self-similarity. This means that the solution resembles itself at many different length scales according to

$$\phi(r, t) = \phi(e^{n\Delta}r, e^{n\Delta}t) \tag{2}$$

where n is an integer and $\Delta \approx 3.44$. Finally, it was found that the critical solutions are also universal in the sense that γ in (1) and Δ in (2) are the same for all one-parameter families of initial data investigated and for any choice of the parameter p . The fact that these three remarkable properties arise as the spacetime is tuned toward the “critical” value p^* justify their designation as *critical phenomena*.

In this paper, I will explore these remarkable phenomena in detail and describe the historical context of their discovery. I will also discuss the analogy between critical phenomena in gravitational collapse and the critical phenomena that arise in second-order phase transitions in statistical mechanics.

2 Mathematical Overview

In his original paper [4], Choptuik studied the gravitational collapse of a massless scalar field coupled to the gravitational field in spherical symmetry. The Einstein equations for a massless scalar field are

$$G_{\mu\nu} = 8\pi \left(\nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} g_{\mu\nu} \nabla_\xi \phi \nabla^\xi \phi \right). \quad (3)$$

The line element is given in spherical-polar coordinates by

$$ds^2 = -\alpha^2(r, t) dt^2 + a^2(r, t) dr^2 + r^2 d\Omega^2 \quad (4)$$

where $a(r, t)$ is the radial metric function, $d\Omega^2$ is the metric of the 2-sphere, and $\alpha(r, t)$ is the lapse function as defined in the ADM formulation of general relativity. The line element (4) can be cast in a more illuminating form by defining the mass-aspect function as $m(r, t) = \frac{1}{2} r (1 - a(r, t)^{-2})$. With this definition we get

$$ds^2 = -\alpha^2(r, t) dt^2 + \left(1 - \frac{2m(r, t)}{r} \right)^{-1} dr^2 + r^2 d\Omega^2 \quad (5)$$

which resembles the Schwarzschild geometry with a time-dependent mass. Note that r can be interpreted as an areal coordinate (that is, it represents the points occupied by origin-centric 2-spheres with surface area $4\pi r^2$). Choptuik employed the polar hypersurface condition [5] under which the extrinsic curvature has only one non-zero component

$$K_r^r = \text{Tr } K. \quad (6)$$

The advantage of this condition is that it forces the radial component of the shift vector β^i to vanish. Moreover, it yields a relatively simple form for the Einstein equations¹,

$$0 = \frac{1}{a} \frac{da}{dr} + \frac{a^2 - 1}{2r} - 2\pi r \left((\partial_r \phi)^2 + \left(\frac{a}{\alpha} \partial_t \phi \right)^2 \right) \quad (7)$$

$$0 = \frac{1}{\alpha} \frac{d\alpha}{dr} - \frac{1}{\alpha} \frac{da}{dr} + \frac{1 - a^2}{r}. \quad (8)$$

The scalar field $\phi(r, t)$ also obeys a wave equation

$$\nabla^\mu \nabla_\mu \phi = \nabla^\mu \partial_\mu \phi = 0. \quad (9)$$

¹For a thorough derivation of the evolution equations, please refer to [6].

In order to solve the evolution equations numerically, it is useful to define new variables that cast the system in first-order form. These variables are

$$\Phi(r, t) = \partial_r \phi(r, t), \quad \Pi(r, t) = \frac{a}{\alpha} \partial_t \phi(r, t). \quad (10)$$

Using (10), the system of equations (7)-(9) becomes

$$\partial_t \Phi = \partial_r \left(\frac{\alpha}{a} \Pi \right) \quad (11)$$

$$\partial_t \Pi = \frac{1}{r^2} \partial_r \left(r^2 \frac{\alpha}{a} \Phi \right) \quad (12)$$

$$0 = \frac{1}{a} \frac{da}{dr} + \frac{a^2 - 1}{2r} - 2\pi r (\Phi^2 + \Pi^2) \quad (13)$$

$$0 = \frac{1}{\alpha} \frac{d\alpha}{dr} - \frac{1}{\alpha} \frac{da}{dr} + \frac{1 - a^2}{r}. \quad (14)$$

These equations are sufficient to numerically evolve the coupled Einstein-scalar field. In order to do so, Choptuik used an $O(h^2)$ finite-difference discretization and the Berger-Oliger algorithm for adaptive mesh refinement. The details of the numerical methods are discussed in [4].

It is notable that the system of equations (11)-(14) are invariant under rescalings $(r, t) \rightarrow (kr, kt)$, as are its solutions. This hints that there is no intrinsic length scale in the problem. It is therefore convenient to re-express the scalar field in terms of form-invariant variables such as

$$X(r, t) \equiv \sqrt{2\pi} \left(\frac{r}{a} \right) \Phi = \sqrt{2\pi} \left(\frac{r}{a} \right) \partial_r \phi \quad (15)$$

$$Y(r, t) \equiv \sqrt{2\pi} \left(\frac{r}{a} \right) \Pi = \sqrt{2\pi} \left(\frac{r}{\alpha} \right) \partial_t \phi. \quad (16)$$

Another advantage of these variables is that the mass-aspect function can be expressed in terms of them:

$$\frac{dm}{dr} = X(r, t)^2 + Y(r, t)^2. \quad (17)$$

This permits us to get a measure of the total mass of a spacetime using

$$M = \int_0^\infty \frac{dm}{dr} dr = \int_0^\infty (X(r, t)^2 + Y(r, t)^2) dr \quad (18)$$

where M is known as the ADM mass.

Finally, solutions to the evolution equations are generated by defining an initial profile $\phi(r, t = 0)$ for the scalar field. The canonical example is an ingoing pulse of the form

$$\phi(r, t = 0; \phi_0, r_0, \Delta, q) = \phi_0 r^3 \exp(-((r - r_0)/\Delta)^q) \quad (19)$$

or equivalently specifying $\Phi(r, t = 0)$ or $\Pi(r, t = 0)$ as defined in (10). Note that while (19) is perhaps the simplest form of initial data, it is not the only choice. Other families studied in Choptuik's original work include

$$\phi(r) = \phi_0 [r^5 (\exp(1/r) - 1)]^{-1} \quad (20)$$

and

$$\phi(r) = \phi_0 \tanh[(r - r_0)/\delta]. \quad (21)$$

In the case of (19) there are 4 parameters that can be varied: ϕ_0 , r_0 , Δ , and q . Denote one of the 4 parameters by p and hold the rest fixed. For example, p could be the central amplitude of the pulse with the width and position held constant, or the width of the pulse with the position and amplitude held constant, or the position of the pulse with the width and amplitude held constant. In any case the initial data becomes constrained to a 1-parameter family of the form $\phi(r, t = 0; p)$. For certain values of $p < p^*$, where p^* is some critical threshold value, the scalar field disperses to infinity upon evolution via the equations above. For other values $p > p^*$, a black hole forms. The game is then to use a bisection search to numerically find the critical value p^* . Choptuik did this by picking an initial data configuration, evolving it for a sufficiently long time and then monitoring the quantity $2m/r$. If a black hole is formed, this quantity will eventually asymptote to $r/2m = 1$.

3 Some Unexpected Phenomena

In general there is no *a priori* way to determine whether some arbitrary initial data will collapse to a black hole or not. Only by evolving many initial data configurations with various values of p can the behaviour near the critical value p^* be uncovered. Choptuik did so by systematically fine-tuning p closer and closer to p^* . This was carried out to the limits of machine precision which at that time was roughly one part in 10^{15} . In the regime of $p \approx p^*$ a plethora of unexpected non-linear behaviour emerged: scale periodicity, power-law mass scaling, and universality.

3.1 Scale Periodicity

One interesting feature of the critical solution is that it exhibits symmetry under changes of scale. Intuitively this means that the solution will resemble itself if you “zoom in” to a smaller length scale or “fast-forward” the solution in time. This can be seen by rewriting the critical solution $\phi(r, t)$ (or equivalently a or α) in terms of new scale-invariant variables

$$x = -\frac{r}{t - t^*}, \quad \tau = -\ln\left(-\frac{t - t^*}{L}\right) \quad (22)$$

where $t < t^*$, and t^* and L depend on the family of initial data under study. When expressed in these coordinates it is found that the critical solution $\phi^*(x, \tau)$ is periodic in τ :

$$\phi^*(x, \tau) = \phi^*(x, \tau + \Delta). \quad (23)$$

In the original (r, t) coordinates this takes the form $\phi(r, t) = \phi(e^{n\Delta}r, e^{n\Delta}t)$ and the solution is said to be scale periodic. In the case of a massless scalar field the value of Δ is found to be $\Delta \approx 3.4$. An illustration of this behaviour is given in Figure 1.

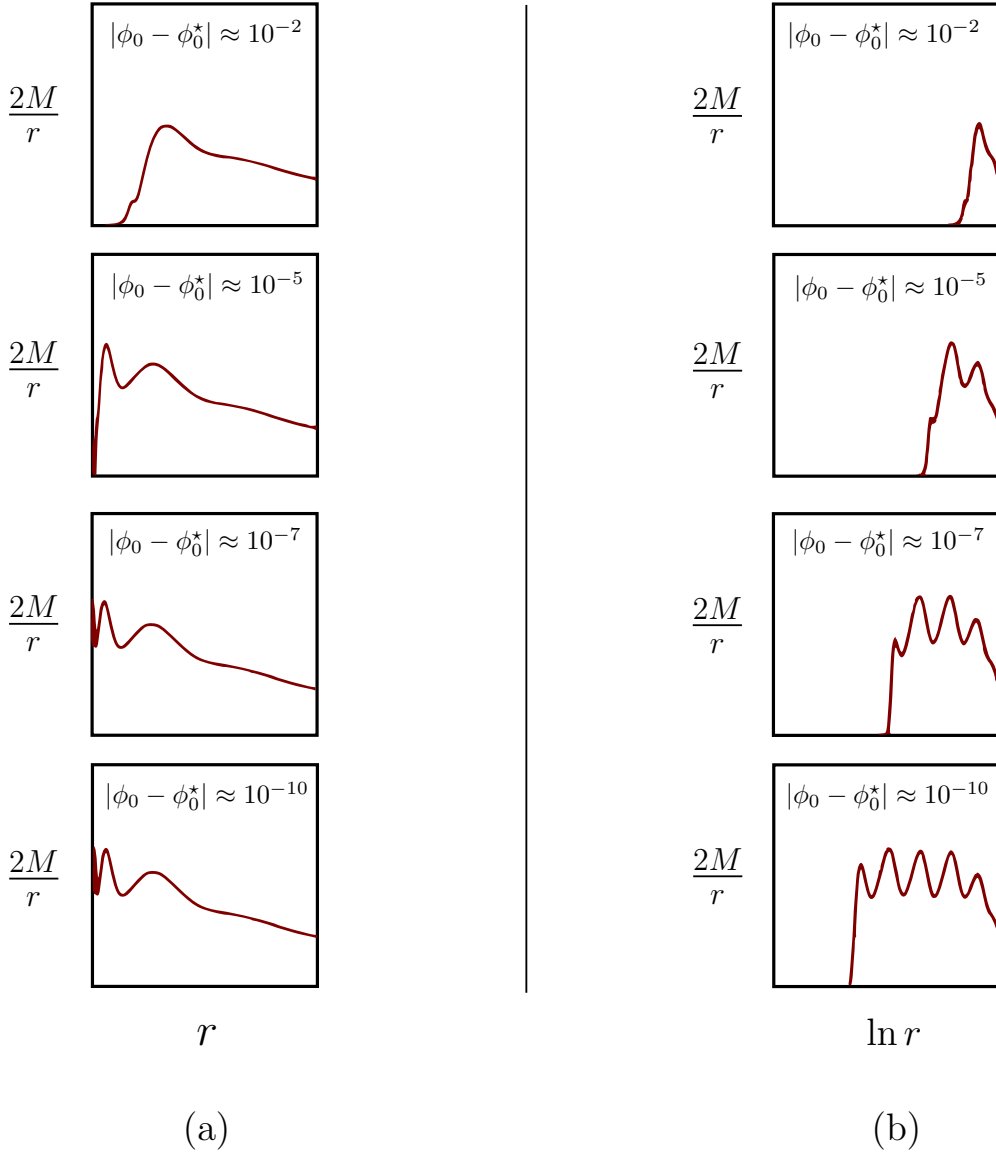


Figure 1: The emergence of scale periodicity in the quantity $2M/r$. Plotted are late-time snapshots from the evolution of four *separate* families of initial data that differed only in their initial scalar field amplitude ϕ_0 . (a) As the amplitude ϕ_0 is tuned closer to the critical value ϕ_0^* (toward the bottom of the figure), oscillations appear for small r . (b) The same data as in (a) except plotted with a logarithmic radial coordinate. In these coordinates we observe the emergence of scale-periodic “echoes” as $\phi_0 \rightarrow \phi_0^*$. If the initial data was exactly critical (i.e. $\phi_0 = \phi_0^*$) then we would expect an infinite number of echoes, each a factor of e^Δ smaller than the previous one. Figure adapted from [3].

3.2 Mass Scaling

By independently varying r_0 , Δ , ϕ_0 and q of (19) away from the critical values, it is found that the black hole mass obeys the scaling relation

$$M \simeq c(p - p^*)^\gamma \tag{24}$$

with the value of γ empirically determined to be $\gamma \approx 0.37$ for any choice of p and any family of initial data². Only the constant c is found to depend on the specific family of initial data that is used.

To understand this result, we can appeal to perturbation theory [7]. The basic idea is to apply linear perturbations around the critical solution ϕ^* (or equivalently a or α) to obtain the scaling exponent γ . Following [8] we expand a near-critical solution ϕ around the critical solution ϕ^* in the coordinates given by (22):

$$\phi(x, \tau; p) = \phi^*(x, \tau; p^*) + \sum_i C_i (p - p^*) e^{-\lambda_i \tau} \xi_i(x, \tau) \tag{25}$$

where ξ_i is an eigenfunction periodic in τ with time dependence $e^{-\lambda_i \tau}$ and C_i are real constants. We assume only one eigenfunction will have $\lambda < 0$ and therefore grow as τ increases [9]. Label this mode by $\lambda_0 < 0$. We therefore expect all other modes with $\lambda > 0$ to eventually decay. This yields

$$\phi(x, \tau_0; p) - \phi^*(x, \tau_0; p^*) = C_0 (p - p^*) e^{-\lambda_0 \tau_0} \xi_0(x, \tau_0) \tag{26}$$

where τ_0 is some late time where all decaying modes have sufficiently died out. Since this is a growing mode we expect it to continue to grow until some later time τ_1 where it becomes of order unity,

$$C_0 (p - p^*) e^{-\lambda_0 \tau_1} \xi_0(x, \tau_1) \approx 1. \tag{27}$$

Taking the logarithm of both sides yields

$$\ln(p - p^*) + \ln(e^{-\lambda_0 \tau_1}) + \text{const.} = 0 \implies \tau_1 \propto \ln(p - p^*)^{1/\lambda_0} \tag{28}$$

Since black hole mass has dimensions of length, and since τ_1 is the only length scale in the above solution, we must have $M \propto \tau_1$. Comparing to (24), this implies $\gamma = 1/\lambda_0$. This is the origin of the mass-scaling relation from dimensional analysis and perturbation theory.

3.3 Universality

The third phenomenon is universality. By this we mean that the critical solution in the limit $p \rightarrow p^*$ is the same for any choice of the parameter p . A second justification for the label “universal” is that the critical phenomena appear for a large class of initial data configurations. This means that no matter whether we take (19), (20) or (21) as the initial scalar field profile, the critical values γ and Δ are found to be the same.

²It is interesting that the value $\gamma \approx 0.37$ is close to $e^{-1} \approx 0.367$. It is unknown whether this is merely a numerical coincidence or if it could be expected from theoretical arguments.

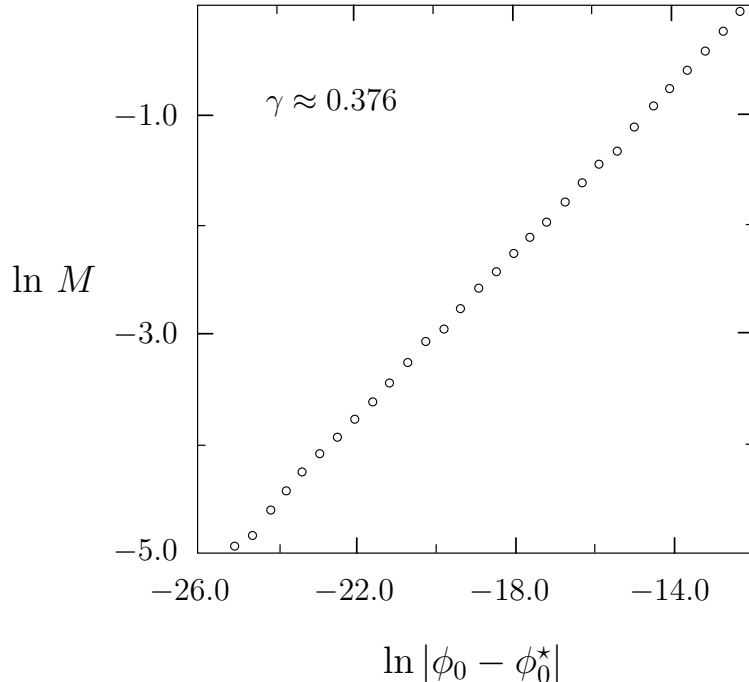


Figure 2: An illustration of the mass-scaling relationship (24) in logarithmic coordinates. It is apparent from the figure that black holes can form with infinitesimal mass in the limit $p \rightarrow p^*$. The value of γ in (24) is found from the slope of this graph to be $\gamma \approx 0.376$. This data is taken from [3].

Universality can be better understood in the context of dynamical systems. Loosely speaking, a dynamical system consists of (1) a phase space in which all possible states of the system are represented and (2) some notion of evolution in time³. For massless scalar fields the phase space is the space of all possible initial data and the time evolution is governed by (11)-(14). The phase space is divided into two *basins of attraction*, which are regions in phase space toward which the system tends to evolve for a wide range of starting points. In our case there are two such basins: one corresponding to collapse to a black hole and the other to Minkowski space (“dispersion to infinity”). There is a natural boundary between these two basins called the *critical surface*. Whether a given set of initial data evolves to a black hole or to Minkowski space depends on which side of the critical surface it is on. Within the critical surface, there exists a third attractor known as the *critical solution*. This is illustrated in Figure 3.

The trajectory of a spacetime in phase space will tend toward one of the attractors. For example, a spacetime that lies in the critical surface (i.e. one that has $p = p^*$) will evolve into the critical solution. A trajectory that lies near the critical surface, but not in it, will first evolve toward the critical spacetime on a trajectory that lies parallel to the critical surface. The parallel movement will slow down when the spacetime is closest to the critical solution. It will linger here for a finite time before finally moving away and evolving toward the Minkowski attractor or the black hole attractor. This lingering behaviour is

³There are some subtleties here that I am ignoring for the sake of brevity. Refer to [8] for a discussion.

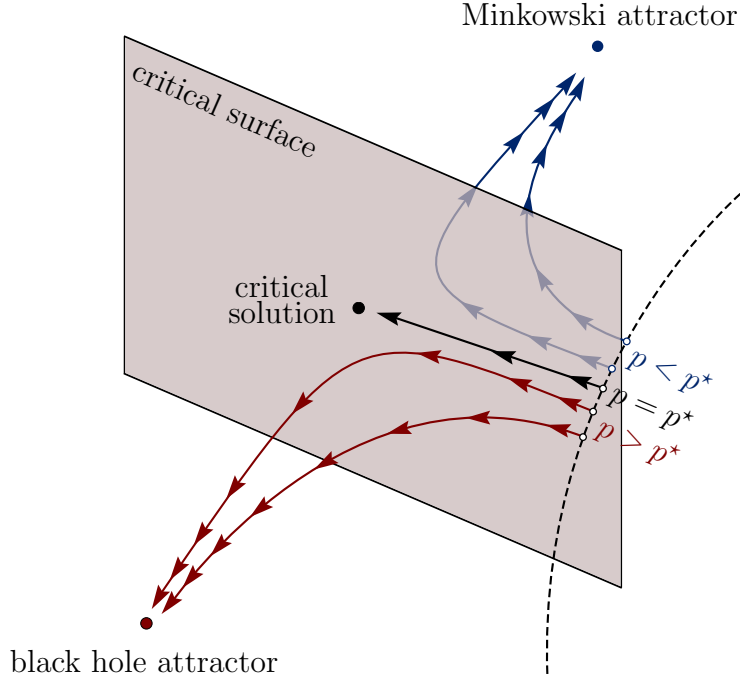


Figure 3: A picture of phase space for the critical collapse of massless scalar fields. The critical surface (black hole threshold) is represented by the grey plane. On each side of the critical surface lies an attractor corresponding to one of the possible end states of the system. The solid lines with arrows represent the time evolutions (trajectories) of solutions in phase space. The density of arrows represent the “speed” of evolution through phase space. The dashed line represents a family of initial data for some parameter p . Note the phenomenon of universality: trajectories that start near the critical surface appear to approach the attractors from the critical solution itself.

another illustration of universality: any near-critical solution will closely resemble the critical solution for some finite time. This means that when a solution finally approaches the black hole attractor it will appear to do so from the critical solution itself. The details of the initial data become irrelevant. The only property that is “remembered” is the original distance from the critical surface, $|p - p^*|$, which manifests in the final mass of the black hole: $M \approx (p - p^*)^\gamma$.

4 Other Matter Models

The discussion in this paper has been limited to massless scalar fields. However, the same critical phenomena will appear in virtually any model that permits black hole formation. The simplest extension is for non-minimal coupling of the scalar radiation [6]. Critical behaviour is found that closely resembles the minimally-coupled case. Another simple model of interest is the massive scalar field [10]. Including a massive term introduces an intrinsic length scale into the problem and therefore breaks scale invariance. With this model, two types of phase transitions emerge depending on the details of the initial data. In “Type I” transitions, a finite mass gap appears at the threshold of black hole formation. This occurs when the

radial extent λ of the initial pulse exceeds the Compton wavelength μ^{-1} : $\lambda\mu > 1$. The label “Type I” comes from the resemblance of this phenomenon to first-order phase transitions in statistical mechanics. The second type is labelled “Type II” and is analogous to the case for massless scalar fields with continuous transitions at the black hole threshold. These occur for $\lambda\mu < 1$ and yield values of $\gamma \approx 0.37$ and $\Delta \approx 3.44$, in agreement with the massless case.

Another study of historical interest was the critical collapse of pure gravitational waves [11]. This was the first study to go beyond spherical symmetry by using axisymmetry to model pulses of ingoing gravitational radiation. Type II behaviour was observed with $\gamma \approx 0.37$ and $\Delta \approx 0.6$, indicating that critical phenomena can occur outside of spherical symmetry.

These early results inspired many subsequent studies of critical collapse in general relativity. Studies have been performed on a wide range of matter models: perfect fluids [12], σ -models [13], and even in theories with higher spacetime dimensions [14]. In virtually all cases, some form of critical phenomenon has been found; this suggests that critical phenomena may indeed be a generic feature of gravitational collapse. However, more work is needed, particularly in models without symmetry restrictions.

5 An Analogy to Statistical Mechanics

In statistical mechanics, the properties of a system are described by a partition function that has the general form

$$Z = \sum_i \exp(-E_i/k_B T) \quad (29)$$

where i is the index that labels the microstates in the system, E_i is the energy of each microstate, T is the temperature and k_B is the Boltzmann constant. The probability P_i that the system occupies the microstate i is simply $P_i = \frac{1}{Z} \exp(-E_i/k_B T)$ which implies that the expectation value of a general observable A is given by

$$\langle A \rangle = \sum_i A_i P_i = \frac{1}{Z} \sum_i A_i \exp(-E_i/k_B T). \quad (30)$$

The canonical example of a statistical mechanical system is the square-lattice Ising model of ferromagnetism. In two dimensions, it is also one of the simplest statistical models to permit a phase transition. For instance, the average magnetization $\langle M \rangle$ for an N -spin system was famously calculated by Onsager [15]:

$$\frac{\langle M \rangle}{N} = \begin{cases} \pm (1 - \sinh^{-4}(2J/k_B T))^{1/8}, & \text{if } T < T_c \\ 0, & \text{if } T > T_c \end{cases} \quad (31)$$

where J is the coupling constant of nearest-neighbour pairs. It is apparent from (31) that something strange occurs at the critical temperature $T = T_c$. Below T_c the system is predominantly ferromagnetic as J dictates the interaction and forces spins to co-align. Above T_c the temperature dominates and the system appears disordered at large scales. At the critical temperature T_c the ordering effect of J and the disordering affect of T compete.

T_c marks a *phase transition* where the macroscopic variable $\langle M \rangle$ changes between non-zero magnetization and disorder. As T_c is approached from below, the magnetization obeys the scaling law

$$|\langle M \rangle| \approx (T - T_c)^\alpha. \quad (32)$$

where α is a non-integer power. It so happens that other macroscopic observables, such as heat capacity and susceptibility, obey an analogous scaling law $A(T) \propto (T - T_c)^\delta$ for some exponent δ . Expressions of this type are known as *critical exponents* and arise generically near continuous (second-order) phase transitions in statistical mechanics.

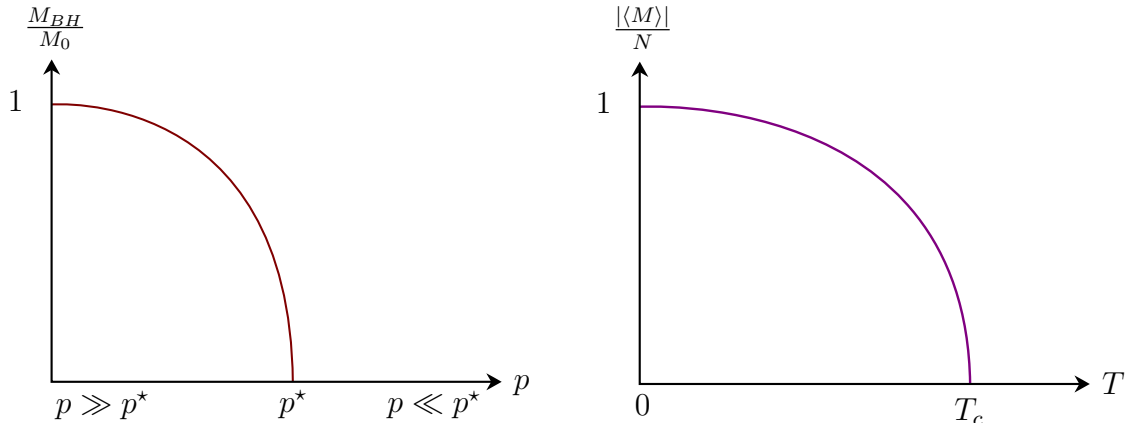


Figure 4: A comparison of the black hole mass scaling law (24) and the 2D Ising magnetization curve (31). The plots are visually identical; it is not surprising that both systems exhibit critical phenomena near their respective critical points, p^* and T_c .

It is striking that the same power-law scaling occurs in statistical mechanics, where the central objects are statistical ensembles of particles, and in general relativity, where the dynamics are described by partial differential equations. This is not a coincidence; it is because these systems share a *renormalization group* [16]. Although a full exposition is beyond the scope of this short report, the basic idea is that the renormalization group reflects the scale-invariant physics that is present in both systems. In the Ising model (for instance), scale-invariance arises at the critical temperature when the correlation length diverges and spins become correlated on infinite length scales. In gravitational collapse, scale-invariance arises at the critical solution when M vanishes and the field variables repeat themselves on periodic length scales. Therefore, one can make the connection that $p - p^*$ plays the role of $T - T_c$, and the mass M of the formed black hole takes the role of the macroscopic observables $\langle A \rangle$ in statistical mechanics.

6 Conclusion

The discovery of critical behaviour at the black hole threshold is perhaps one of the biggest revelations to come from numerical general relativity. Although nearly 30 years have passed since its discovery, the inescapable question is *why* virtually every matter model studied to date has admitted a critical solution. It appears that critical behaviour is generic feature

of gravitational collapse but there is no clear explanation for why this should be so. Some heuristic arguments can be made from a dynamical systems picture, dimensional analysis and perturbation theory but there are still many questions that remain unanswered.

There is still much to be said of critical phenomena that could not be included in this paper. Of particular interest is the potential violation of cosmic censorship through the formation of naked singularities. Another topic only touched on briefly was the difference between Type I and Type II collapse. For comprehensive discussions of these ideas and more, I will refer the reader to an excellent review paper [8].

Acknowledgements

I am grateful to Matthew Choptuik for useful discussions.

References

- [1] D. Christodoulou, *Commun. Math. Phys.* **109**, 613 (1987).
- [2] D. Christodoulou, *Commun. Pure Appl. Math.* **44**, 339 (1991).
- [3] M. W. Choptuik, in *Deterministic Chaos in General Relativity*, edited by D. Hobill, (Springer, Boston, 1994).
- [4] M. W. Choptuik, *Phys. Rev. Lett.* **70**, 9 (1993).
- [5] J. M. Bardeen and T. Piran, *Phys. Rep.* **96**, 205 (1983).
- [6] M. W. Choptuik, in *Approaches to Numerical Relativity*, edited by R. d’Inverno, (Cambridge Univ. Press, Cambridge, 1992).
- [7] T. Koike, T. Hara, and S. Adachi, *Phys. Rev. Lett.* **74**, 5170 (1995).
- [8] C. Gundlach and J. M. Martín-García, *Living Rev. Relativ.* **10**, 5 (2007).
- [9] F. Pretorius, *Numerical simulations of gravitational collapse*, PhD thesis, University of British Columbia, 2002.
- [10] P. R. Brady, C. M. Chambers, and S. M. C. V. Gonçalves, *Phys. Rev. D* **56**, R6057 (1997).
- [11] A. M. Abrahams and C. R. Evans, *Phys. Rev. Lett.* **70**, 2980 (1993).
- [12] C. R. Evans and J. S. Coleman, *Phys. Rev. Lett.* **72**, 1782 (1994).
- [13] S. L. Liebling and M. W. Choptuik, *Phys. Rev. Lett.* **77**, 1424 (1996).
- [14] P. Bizoń, T. Chmaj, A. Rostworowski, B. G. Schmidt, and Z. Tabor, *Phys. Rev. D* **72**, 121502 (2005).

REFERENCES

- [15] L. Onsager, Phys. Rev. **65**, 117 (1944).
- [16] T. Hara, T. Koike, and S. Adachi, [arXiv:gr-qc/9607010] (1996).