FORMATION OF SINGULARITIES
IN YANG–MILLS EQUATIONS*

PIOTR BIZOŃ

M. Smoluchowski Institute of Physics, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland

This is a survey of recent studies of singularity formation in solutions of spherically symmetric Yang–Mills equations in higher dimensions. The main attention is focused on five space dimensions because this case exhibits interesting similarities with Einstein's equations in the physical dimension, in particular the dynamics at the threshold of singularity formation shares many features (such as universality, self-similarity, and scaling) with critical phenomena in gravitational collapse. The borderline case of four space dimensions is also analysed and the formation of singularities is shown to be intimately tied to the existence of the instanton solution.

PACS numbers: 11.10.Lm, 04.20.Dw

1. Introduction

One of the most interesting features of many nonlinear evolution equations is the spontaneous onset of singularities in solutions starting from perfectly smooth initial data. Such a phenomenon, usually called "blowup", has been a subject of intensive studies in many fields ranging from fluid dynamics to general relativity. Whether or not the blowup can occur for a given nonlinear evolution equation is the central mathematical question which, from the physical point of view, has a direct bearing on our understanding of the limits of validity of the corresponding model. Unfortunately, this is often a difficult question. Two famous examples, for which the answer is not known, are the Navier–Stokes equation and the Einstein equations. Once the existence of blowup is established for a particular equation, many

---

further questions come up, such as: When and where does the blowup occur? What is the character of blowup and is it universal? Can a solution be continued past the singularity?

In this paper we consider these questions for the Yang–Mills (YM) equations in higher dimensions. In the physical $3 + 1$ dimensions, where the YM equations are the basic equations of gauge theories describing the weak and strong interactions of elementary particles, it is known that no singularities can form. This was shown by Eardley and Moncrief [1] who proved that solutions starting from smooth initial data remain smooth for all future times. The motivation for studying the YM equations in unphysical $D + 1$ dimensions for $D > 3$ is twofold (and unrelated to the latest fashion of doing physics in extra dimensions). From the mathematical point of view, it is the obvious thing to ask how the property of global regularity depends on the dimension of the underlying spacetime and whether singularities can form in $D + 1$ dimensions for $D > 3$. However, there is also a less evident physical reason which is motivated by the hope that by understanding the problem of singularity formation for the YM equations one might get insight into the analogous, but much more difficult, problem in general relativity. From this viewpoint — in which the YM equations are considered as a toy model for the Einstein equations — it is essential that these two equations belong to the same criticality class. Let us recall that the criticality class is defined as the degree α in the homogeneous scaling of energy $E \to \lambda^\alpha E$ under dilations $x \to x/\lambda$. The classification of equations into subcritical ($\alpha < 0$), critical ($\alpha = 0$), and supercritical ($\alpha > 0$) is a basis of the heuristic meta-principle according to which subcritical equations are globally regular, while supercritical equations may develop singularities for some (large) initial data [2]. For the YM equations we have $\alpha_{YM} = D - 4$, while for the Einstein equations $\alpha_E = D - 2$. Therefore, the YM equations in $D = 5$ have the same criticality, $\alpha = 1$, as the Einstein equations in the physical dimension. Another way of seeing this is to note that in $D = 5$ the dimension of the YM coupling constant $[e^2] = M^{-1}L^{D-4}$ (in $c = 1$ units) is the same as the dimension of the physical Newton's constant $[G] = M^{-1}L$.

For the reason just explained, the main body of this paper is focused on the lowest super-critical dimension $D = 5$. In Section 3 we show that in this case there exists a countable family of regular (by regularity we mean analyticity inside the future light cone) spherically symmetric self-similar solutions labelled by a nonnegative integer n (a nodal number). Next, using linear stability analysis we show in Section 4 that the number of unstable modes around a given solution is equal to its nodal number. The role of self-similar solutions in the dynamical evolution is studied in Section 5, where we show that: *(i)* the $n = 0$ solution determines a universal asymptotics of singularity formation for solutions starting from generic "large" initial

data; *(ii)* the $n = 1$ solution plays the role of a critical solution sitting at the threshold of singularity formation. The latter is in many respects similar to the critical behaviour at the threshold of black hole formation in gravitational collapse. In both cases the threshold of singularity (or black hole) formation can be identified with the codimension-one stable manifold of a self-similar solution with exactly one unstable mode. These similarities are discussed in detail in Section 6.

We consider also the Cauchy problem for the YM equations in $D = 4$. Despite intensive studies of this borderline case, the problem of global existence is open. In Section 7 we describe numerical simulations which, in combination with analytic results, strongly suggest that large-energy solutions do blow up. We show that the process of singularity formation is due to concentration of energy and proceeds via adiabatic shrinking of the instanton solution. At the end, a recent attempt of determining the asymptotic rate of shrinking is sketched.

We remark that there are close parallels between YM equations in $D + 1$ dimensions and wave maps in $(D - 2) + 1$ dimensions [3]. Indeed, many of the phenomena described here are mirrored for the equivariant wave maps into spheres in three [4,5] and two [6] spatial dimensions.

Sections 5 and 7 of this survey are based on joint work with Tabor [7]. The material of Sections 3 and 4 is new.

2. Setup

We consider Yang–Mills (YM) fields in $D + 1$ dimensional Minkowski spacetime (in the following Latin and Greek indices take the values $1, 2, \ldots, D$ and $0, 1, 2, \ldots, D$ respectively). The gauge potential A_α is a one-form with values in the Lie algebra g of a compact Lie group G. In terms of the curvature $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta]$ the action is

$$S = \frac{1}{e^2} \int \mathrm{Tr}\, (F_{\alpha\beta} F^{\alpha\beta}) d^D x \, dt \,, \tag{1}$$

where e is the gauge coupling constant. Hereafter we set $e = 1$. The YM equations derived from (1) are

$$\partial_\alpha F^{\alpha\beta} + [A_\alpha, F^{\alpha\beta}] = 0 \,. \tag{2}$$

As written, this equation is underdetermined because of the gauge invariance

$$A_\alpha \to U^{-1} A_\alpha U + U^{-1} \partial_\alpha U \,, \tag{3}$$

where U is an arbitrary function with values in G. In order to correctly formulate the Cauchy problem for equation (2), one must impose additional

conditions which fix this gauge ambiguity. We shall not discuss this issue here because in the spherically symmetric ansatz, to which this paper is restricted, the gauge is fixed automatically.

For simplicity, we take here $G = SO(D)$ so the elements of $SO(D)$ can be considered as skew-symmetric $D \times D$ matrices and the Lie bracket is the usual commutator. Assuming the spherically symmetric ansatz [8]

$$A_\mu^{ij}(x) = \left(\delta_\mu^i x^j - \delta_\mu^j x^i\right) \frac{1 - w(t, r)}{r^2}, \tag{4}$$

the YM equations reduce to the scalar semilinear wave equation for the magnetic potential $w(t, r)$

$$-w_{tt} + \Delta_{(D-2)} w + \frac{D-2}{r^2} w(1 - w^2) = 0, \tag{5}$$

where $\Delta_{(D-2)} = \partial_r^2 + \frac{D-3}{r}\partial_r$ is the radial Laplacian in $D - 2$ dimensions. The central question for equation (5) is: can solutions starting from smooth initial data

$$w(0, r) = f(r), \qquad w_t(0, r) = g(r) \tag{6}$$

become singular in future? As mentioned above, in the physical $D = 3$ dimensions Eardley and Moncrief answered this question in the negative [1]. However, simple heuristic arguments indicate that the property of global regularity enjoyed by the YM equations in $D = 3$ might break down in higher dimensions. In order to see why the global behaviour of solutions is expected to depend critically on the dimension D, we recall two basic facts. The first fact is the conservation of (positive definite) energy

$$E = \int\limits_{R^D} \mathrm{Tr}\, \left(F_{0i}^2 + F_{ij}^2\right) d^D x = c(D) \int\limits_0^\infty \left(w_t^2 + w_r^2 + \frac{D-2}{2r^2}(1 - w^2)^2\right) r^{D-3} dr, \tag{7}$$

where the coefficient $c(D) = (D-1)\mathrm{vol}(S^{D-1})$ follows from the integration over the angles and taking the trace. The second fact is scale-invariance of the YM equations: if $A_\alpha(x)$ is a solution of (2), so is $\tilde{A}_\alpha(x) = \lambda^{-1} A_\alpha(x/\lambda)$, or equivalently, if $w(t, r)$ is a solution of (5), so is $\tilde{w}(t, r) = w(t/\lambda, r/\lambda)$. Under this scaling the energy scales as $\tilde{E} = \lambda^{D-4} E$, hence the YM equations are subcritical for $D \leq 3$, critical for $D = 4$, and supercritical for $D \geq 5$. In the subcritical case, shrinking of solutions to arbitrarily small scales costs infinite amount of energy, so it is forbidden by energy conservation. In other words, transfer of energy to arbitrarily high frequencies is impossible and consequently the Cauchy problem should be well posed in the energy norm. This important fact was proved in $D = 3$ by Klainerman and Machedon [9],

who thereby strengthened the result of Eardley and Moncrief. In the supercritical case, shrinking of solutions might be energetically favourable and consequently singularities are anticipated. In fact, we shall show below that singularities do form already in the lowest supercritical dimension $D = 5$. In the critical dimension $D = 4$ the problem of singularity formation is more subtle because the scaling argument is inconclusive.

3. Self-similar solutions in $D = 5$

In order to set the stage for the discussion of singularity formation we first need to analyse in detail the structure of self-similar solutions of equation (5). As we shall see, these solutions play a key role in understanding the nature of blowup. By definition, self-similar solutions are invariant under dilations $w(t, r) \to w(t/\lambda, r/\lambda)$, hence they have the form

$$w(t, r) = W(\eta), \qquad \eta = \frac{r}{T - t}, \tag{8}$$

where a positive constant T, clearly allowed by the time translation invariance, is introduced for later convenience. Note that for a self-similar solution we have

$$\partial_r^2 W(\eta)\Big|_{r=0} = \frac{1}{(T - t)^2} W''(0), \tag{9}$$

hence the solution becomes singular at the centre when $t \to T$ (there is no blowup in the first derivative because regularity demands that $W'(0) = 0$). Thus, each self-similar solution $W(\eta)$ provides an explicit example of a singularity developing in finite time from smooth initial data.

Substituting the ansatz (8) into (5) one obtains the ordinary differential equation

$$W'' + \left(\frac{D - 3}{\eta} + \frac{(D - 5)\eta}{1 - \eta^2} \right) W' + \frac{D - 2}{\eta^2 (1 - \eta^2)} W(1 - W^2) = 0. \tag{10}$$

As explained in the introduction, because of the expected connections with the Einstein equations, we are mainly interested in the lowest super-critical dimension $D = 5$. In this case equation (10) reduces to

$$W'' + \frac{2}{\eta} W' + \frac{3}{\eta^2 (1 - \eta^2)} W(1 - W^2) = 0. \tag{11}$$

Although the similarity coordinate η is natural in the discussion of singularity formation, it has a disadvantage of not covering the region $t > T$, in particular it does not extend to the future light cone of the point $(T, 0)$. For this reason we define a new coordinate $x = 1/\eta$ which covers the whole

spacetime: the past and the future light cones are located at $x = 1$ and $x = -1$, respectively; while the centre $r = 0$ corresponds to $x = \infty$ (for $t < T$) and $x = -\infty$ (for $t > T$). In terms of x equation (11) becomes

$$(x^2 - 1)W'' + 3W(1 - W^2) = 0 \,. \tag{12}$$

We first consider this equation inside the past light cone, that is for $1 \leq x < \infty$ and impose the boundary conditions

$$W(1) = 0 \quad \text{and} \quad W(\infty) = \pm 1 \,, \tag{13}$$

which follow from the demand of smoothness at the endpoints. As we shall see below, once a solution to this boundary value problem is constructed, its extension beyond the past light cone can be easily done.

To show that equation (12) admits solutions satisfying (13) we shall employ a shooting technique. The main idea of this method is to replace the boundary value problem by the initial value problem with initial data imposed at one of the endpoints and then adjusting these data so that the solution hits the desired boundary value at the second endpoint. In the case at hand we shall shoot from $x = 1$ towards infinity. Substituting a formal power series expansion about $x = 1$ into (12) one finds the asymptotic behaviour

$$W(x) = a(x - 1) - \frac{3a}{4}(x - 1)^2 + O\left((x - 1)^3\right) \,, \tag{14}$$

where a is a free parameter determining uniquely the whole series. In the following a solution of equation (12) starting at $x = 1$ with the asymptotic behaviour (14) will be called an a-orbit. Without loss of generality we may assume that $a \geq 0$. We claim that there is a countable set of values $\{a_n\}$ for which the a_n-orbits exist for all $x \geq 1$ and have the desired asymptotics at infinity (such orbits will be called connecting). The proof consists of several steps.

Step 1 (Local existence). First, we need to show that a-orbits do in fact exist, that is, the series (14) has a nonzero radius of convergence. Since the point $x = 1$ is singular, this fact does not follow from standard theorems. Fortunately, in [10] Breitelohner, Forgács, and Maison have derived the following result concerning the behaviour of solutions of a system of ordinary differential equations near a singular point:

Theorem [BFM]. Consider a system of first order differential equations for $n + m$ functions $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_m)$

$$y\frac{du_i}{dy} = y^{\mu_i} f_i(y, u, v), \qquad y\frac{dv_i}{dy} = -\lambda_i v_i + y^{\nu_i} g_i(y, u, v), \tag{15}$$

where constants $\lambda_i > 0$ and integers $\mu_i, \nu_i \geq 1$ and let C be an open subset of R^n such that the functions f and g are analytic in the neighbourhood of $y = 0, u = c, v = 0$ for all $c \in C$. Then there exists an n-parameter family of solutions of the system (15) such that

$$u_i(y) = c_i + O(y^{\mu_i}), \qquad v_i(y) = O(y^{\nu_i}), \tag{16}$$

where $u_i(y)$ and $v_i(y)$ are defined for all $c \in C, |y| < y_0(c)$ and are analytic in y and c.

We shall make use of this theorem to prove the local existence of a-orbits. In order to put equation (12) into the form (16) we define the variables

$$y = x - 1, \quad u(y) = W', \quad v(y) = \frac{W}{x-1} - W', \tag{17}$$

and get

$$yv' = -v + yf, \qquad yu' = yf, \qquad f = \frac{3(u+v)\left[1 - y^2(u+v)^2\right]}{2+y}. \tag{18}$$

Since the function $f(y, u, v)$ is analytic near $y = 0$ for any u and v, according to the BFM theorem, there exists a one-parameter family of local solutions such that

$$u(y) = a + O(y), \qquad v(y) = O(y). \tag{19}$$

Transforming (19) back to the original variables we obtain the behaviour of a-orbits.

Step 2 (A priori global behaviour). It follows immediately from (12) that for $x > 1$ a solution cannot have a maximum (resp. minimum) for $W > 1$ (resp. $W < -1$). Thus, once the solution leaves the strip $|W| < 1$, it cannot reenter it (actually, such a solution becomes singular for a finite x). It is also clear that as long as $|W| < 1$ the solution cannot go singular. To derive the asymptotics at infinity of a-orbits that stay in the strip $|W| < 1$ we shall make use of the following functional

$$Q(x) = \frac{1}{2}(x^2 - 1)W'^2 - \frac{3}{4}(1 - W^2)^2. \tag{20}$$

For solutions of equation (12) we have

$$Q'(x) = xW'^2, \tag{21}$$

so $Q(x)$ is monotone increasing. Now, we shall show that solutions satisfying $|W| < 1$ for all $x \geq 1$ tend to $W = \pm 1$ as $x \to \infty$. To see this, first notice

that for such solutions Q must be negative because if $Q(x_0) > 0$ for some $x_0 > 1$ then $|W'|$ is strictly positive for $x > x_0$ so the solution must leave the strip $|W| < 1$ in finite time. Since $Q' \geq 0$ and $Q \leq 0$, it follows that Q has a nonpositive limit at infinity which in turn implies by (21) that $\lim_{x \to \infty} xW' = 0$ and by (20) that $\lim_{x \to \infty} W$ exists. By L'Hôpital's rule we have $\lim_{x \to \infty} x^2 W'' = 0$ and using (12) again, we get that $\lim_{x \to \infty} W$ equals ± 1 or 0. The latter is impossible because then $Q(\infty) = -3/4$ is a global minimum contradicting the fact that Q increases. Thus, $W(\infty) = \pm 1$.

Step 3 (i) (Behaviour of a-orbits for small a). Rescaling $w(x) = W(x)/a$ we get

$$(x^2 - 1)w'' + 3w(1 - a^2 w^2) = 0, \qquad w(1) = 0, \quad w'(1) = 1. \qquad (22)$$

As $a \to 0$, the solutions of this equation tend uniformly on compact intervals to the solution of the limiting equation

$$(x^2 - 1)w'' + 3w = 0 \qquad (23)$$

with the same initial condition. This equation can be solved explicitly but for the purpose of the argument it suffices to notice that its solution, call it $w_{\mathrm{L}}(x)$, is oscillating at infinity. Since $W(x, a) \approx a w_{\mathrm{L}}(x)$ up to an arbitrarily large x if a is sufficiently small, it follows that the number of zeros of the solution $W(x, a)$ tends to infinity as $a \to 0$.

(ii) (Behaviour of a-orbits for large a.) We rescale the variables, setting $y = a(x - 1)$, $\bar{w}(y) = W(x)$ to get

$$\bar{y}(y + 2a)\bar{w}'' + 3\bar{w}(1 - \bar{w}^2) = 0, \qquad \bar{w}(0) = 0, \quad \bar{w}'(0) = 1. \qquad (24)$$

As $a \to \infty$, the solutions of this equation tend uniformly on compact intervals to the solution $\bar{w}(y) = y$ of the limiting equation $\bar{w}'' = 0$. Thus, $W(x, a) \approx a(x - 1)$ for large a and therefore the a-orbit crosses $W = 1$ for a finite x.

Step 4 (Shooting argument). We define the set

$$A_0 = \{a \mid W(x, a) \text{ strictly increases up to some } x_0 \text{ where } W(x_0, a) = 1\}. \qquad (25)$$

We know from Step 3 that the set A_0 is nonempty (because the a-orbits with large a belong to it) and bounded below (because the a-orbits with small a do not belong to it). Thus $a_0 = \inf A_0$ exists. The solution $W(x, a_0)$ cannot cross the line $W = 1$ at a finite x because the same would be true for nearby solutions, violating the definition of a_0. Thus, $0 \leq W(x, a_0) < 1$ for all x and hence, by Step 2, $\lim_{x \to \infty} W(x, a_0) = 1$. This completes the proof of existence of the nodeless self-similar solution $W_0(x) \stackrel{\mathrm{def}}{=} W(x, a_0)$.

Next, let us consider the solution with $a = a_0 - \varepsilon$ for small $\varepsilon > 0$. By the definition of a_0 there must be a point x_0 where this solution attains a positive local maximum $W(x_0) < 1$ and since no minima are possible for $0 < W < 1$, it follows that there must be a point $x_1 > x_0$ where $W(x_1, b) = 0$. We shall show that $Q(x_1, a) > 0$ provided that ε is sufficiently small. As argued above this would imply that the solution $W(x, a)$ leaves the strip $|W| < 1$ via $W = -1$. From (21) we have

$$Q(x_1) - Q(x_0) = \int\limits_{x_0}^{x_1} x W'^2 dx = - \int\limits_{0}^{W(x_0)} x W' dW. \tag{26}$$

In order to estimate the last integral note that for $x > x_0$

$$Q(x) - Q(x_0) = \frac{1}{2}(x^2 - 1)W'^2 - \frac{3}{4}(1 - W^2)^2 + \frac{3}{4}(1 - W^2(x_0))^2 > 0, \tag{27}$$

so $x|W'| > \sqrt{\frac{3}{2}} \sqrt{(1 - W^2)^2 - (1 - W^2(x_0))^2}$. Substituting this into (26) one gets

$$Q(x_1) > -\frac{3}{4}\left(1 - W^2(x_0)\right)^2 + \sqrt{\frac{3}{2}} \int\limits_{0}^{W(x_0)} \sqrt{(1 - W^2)^2 - (1 - W^2(x_0))^2} \, dW. \tag{28}$$

The right hand side of this inequality is equal to $\sqrt{2/3}$ for $W(x_0) = 1$ so, by continuity, it remains strictly positive for $W(x_0)$ near 1. By taking a sufficiently small ε we can have $W(x_0)$ arbitrarily close to 1, hence $Q(x_1) > 0$ which proves that a-orbits with $a = a_0 - \varepsilon$ have exactly one zero. This means that the set $A_1 = \{a \,|\, W(x, a)$ increases up to some x_0 where it attains a positive local maximum $W(x_0) < 1$ and then decreases monotonically up to some x_1 where $W(x_1) = -1\}$ is nonempty. Let $a_1 = \inf A_1$. By Step 3, a_1 exists and is strictly positive. Using the same argument as above we conclude that the a_1-orbit must stay in the region $|W| < 1$ for all x, hence $\lim_{x \to \infty} W(x, a_1) = -1$. This completes the proof of existence of the self-similar solution $W_1(x) \stackrel{\text{def}}{=} W(x, a_1)$ with exactly one zero.

The subsequent connecting orbits are obtained by induction. We conclude that there exists a countable family of self-similar solutions $W_n(x)$ indexed by the integer $n = 0, 1, \ldots$ n which counts the number of zeros for $x > 1$.

Remark. Since the sequence $\{a_n\}$ is decreasing and bounded below by zero, it has a nonnegative limit $\lim_{n \to \infty} a_n = a^* \geq 0$. If $a^* > 0$, then the a^*-orbit cannot leave the region $|W| < 1$ for a finite x (because the set of such orbits

is clearly open) hence it must be a connecting orbit with some finite number of zeros. But this contradicts the fact that the number of zeros of a_n-orbits increases with n. Hence, $a^* = 0$. This implies that for any finite x, $W_n(x)$ goes to zero when $n \to \infty$.

We remark that the existence of the solution W_0 was first shown by Cazenave, Shatah, and Tahvildar-Zadeh [3] via a variational method.

The shooting technique is not only a powerful analytical tool; it is also an efficient numerical method of solving two-point boundary value problems. The numerical results produced by this method are shown in Table I and figure 1.

TABLE I

The shooting parameters of solutions W_n for $n \leq 5$.

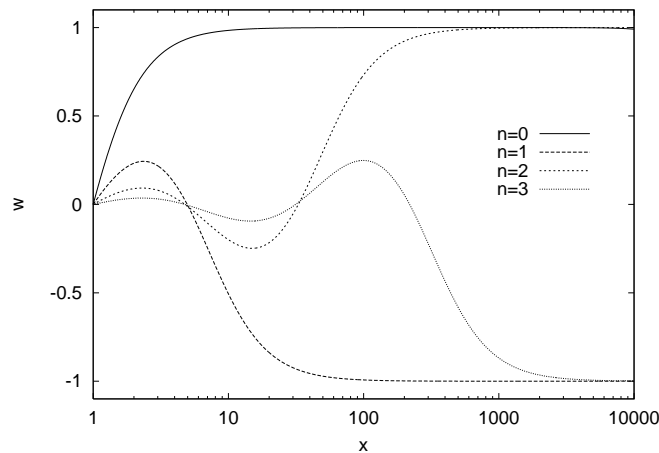| n	0	1	2	3	4	5
a_n	1.25	0.4813158	0.1864517	0.0722966	0.02803703	0.01087315

Fig. 1. The first four self-similar solutions $W_n(x)$.

Surprisingly, it turned out that $a_0 = 5/4$ (with very good accuracy). This was a hint that the solution W_0 has a simple closed form. Indeed, playing with the power series expansion (14) we found that

$$W_0(x) = \frac{x^2 - 1}{x^2 + \frac{3}{5}}. \tag{29}$$

Below we show an amusing calculation by Maple which helped us in finding this formula.

```
>   restart;
>   with(DEtools):
>   with(numapprox):
>   ode:=(x^2-1)*diff(w(x),x$ 2)+3*w(x)*(1-w(x)^2)=0;
```

$$ode := (x^2 - 1)\left(\tfrac{\partial^2}{\partial x^2}\, w(x)\right) + 3\, w(x)\,(1 - w(x)^2) = 0$$

```
>   ic:=w(1)=0,D(w)(1)=5/4;
```

$$ic := w(1) = 0,\; D(w)(1) = \frac{5}{4}$$

```
>   sol:=dsolve({ode,ic},w(x));
```

$$sol :=$$

```
>   sol_formal:=rhs(dsolve({ode,ic},w(x),type=series));
```

$$sol\_formal :=$$

$$\frac{5}{4}\,(x-1) - \frac{15}{16}\,(x-1)^2 + \frac{25}{64}\,(x-1)^3 + \frac{25}{256}\,(x-1)^4 - \frac{375}{1024}\,(x-1)^5 + O((x-1)^6)$$

```
>   pade_sol:=pade(sol_formal,x=1,[2,2]);
```

$$pade\_sol := \frac{\dfrac{5}{8}\,(x-1)^2 + \dfrac{5}{4}\,x - \dfrac{5}{4}}{-\dfrac{1}{4} + \dfrac{5}{4}\,x + \dfrac{5}{8}\,(x-1)^2}$$

```
>   sol:=simplify(pade_sol);
```

$$sol := 5\,\frac{x^2 - 1}{3 + 5\,x^2}$$

```
>   subs(w(x)=sol,ode);
```

$$(x^2 - 1)\left(\frac{\partial^2}{\partial x^2}\left(5\,\frac{x^2 - 1}{3 + 5\,x^2}\right)\right) + 15\,\frac{(x^2 - 1)\left(1 - 25\,\dfrac{(x^2 - 1)^2}{(3 + 5\,x^2)^2}\right)}{3 + 5\,x^2} = 0$$

```
>   simplify(%);
```

$$0 = 0.$$

So far our analysis of self-similar solutions was restricted to the interior of the past light cone of the singularity. To show that the solutions W_n represent genuine naked singularities, we need to extend them to the future light cone, that is to $x = -1$. Fortunately, such an extension creates no problem because an a-orbit shot backwards from $x = 1$ cannot go singular before reaching $x = -1$. This follows immediately from (12) by observing that, in the interval $-1 < x < 1$, $W(x)$ is concave down (resp. up) for $W > 1$ (resp. $W < -1$), hence $W(x)$ remains bounded as $x \to -1^+$. Moreover, the function $Q(x)$ is negative and decreasing near $x = -1$, thus $\lim_{x \to -1^+} Q(x)$ exists which implies in turn that $c = \lim_{x \to -1^+} W(x)$ exists. Having that, the standard asymptotic analysis gives the following leading order behaviour for $x \to -1^+$

$$W(x) \sim c + \frac{3}{2}c(1 - c^2)(x + 1)\ln(x + 1)\,. \tag{30}$$

The singular logarithmic term in (30) can be eliminated by fine-tuning the shooting parameter a, however this is not expected to happen for the solutions $W_n(x)$ because in their construction the freedom of adjusting a was already used to tune away the singular behaviour for $x > 1$. We conclude that the self-similar solutions W_n are C^0 at the future light cone and are analytic everywhere below it. The only (somewhat surprising) exception is the solution W_0 which is analytic in the entire spacetime.

4. Linear stability of self-similar solutions

In this section we study the linear stability of self-similar solutions W_n. This analysis is essential in determining the role of self-similar solutions in dynamics. We restrict attention to the interior of the past light cone of the point $(T, 0)$ and define the new time coordinate $s = -\ln\sqrt{(T - t)^2 - r^2}$. Note that $s \to \infty$ when $t \to T$, and the lines of constant s are orthogonal to the rays of constant x. In terms of s and x, equation (5) becomes (for $D = 5$)

$$-\frac{e^{2s}}{x^2 - 1}(e^{-2s}w_s)_s + (x^2 - 1)w_{xx} + 3w(1 - w^2) = 0\,. \tag{31}$$

Of course, this equation reduces to (12) if w does not depend on s. In order to determine the stability of self-similar solutions W_n we seek solutions of (31) in the form $w(s, x) = W_n(x) + v(s, x)$. Neglecting the $O(v^2)$ terms we obtain the linear evolution equation for the perturbation $v(s, x)$

$$-\frac{e^{2s}}{x^2 - 1}(e^{-2s}v_s)_s + (x^2 - 1)v_{xx} + 3(1 - 3W_n^2)v = 0\,. \tag{32}$$

Substituting $v(s, x) = e^{(\alpha+1)s} \sqrt{x^2 - 1}\, u(x)$ into (32) we get the eigenvalue problem in the standard Sturm–Liouville form

$$-\frac{d}{dx}\left((x^2 - 1)\frac{du}{dx}\right) - 3(1 - 3W_n^2)u = \frac{\lambda}{x^2 - 1}u\,, \qquad (33)$$

where $\lambda = -\alpha^2$. Using the variable $\rho = \frac{1}{2}\ln(\frac{x-1}{x+1})$ ranging from zero to infinity we transform (33) into the radial Schrödinger equation

$$-\frac{d^2 u}{d\rho^2} + V_n u = \lambda u\,, \qquad V_n = -\frac{3(1 - 3W_n^2)}{\sinh^2\rho}\,. \qquad (34)$$

The potential $V_n(\rho)$ has a typical "quantum mechanical" shape (see figure 2) with the asymptotics

$$V_n(\rho) \sim \begin{cases} 6/\rho^2 & \text{for } \rho \to 0\,, \\ -12\exp(-2\rho) & \text{for } \rho \to \infty\,. \end{cases} \qquad (35)$$

Note that the potential can be expressed in the form $V_n(\rho) = l(l+1)/\rho^2 + V_n^{\text{reg}}(\rho)$ with $l = 2$, where the regular part $V_n^{\text{reg}}(\rho)$ is everywhere negative and $V_n^{\text{reg}}(0) \to -\infty$ as $n \to \infty$.
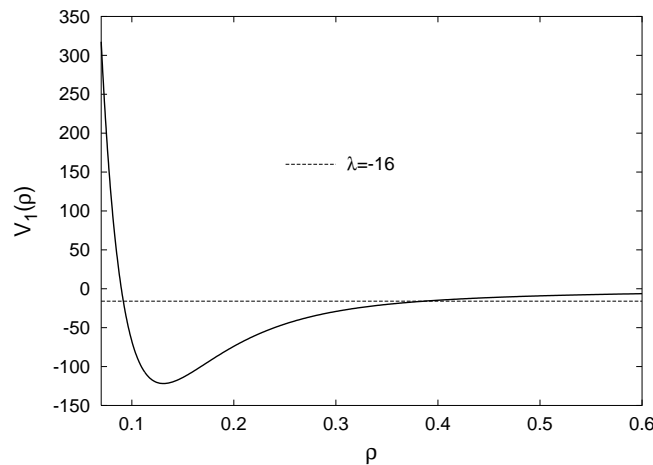

Fig. 2. The potential for the perturbations around the self-similar solution W_1. The single bound state with energy $\lambda = -16$ is indicated.

Both endpoints $\rho = 0$ and $\rho = \infty$ are of the limit-point type, that is, exactly one solution near each point is square-integrable (admissible). Near $\rho = 0$ the admissible solutions behave as $u(\rho) \sim \rho^3$. For $\rho \to \infty$ and $\lambda < 0$ the admissible solutions behave as $u(\rho) \sim e^{-\alpha\rho}$ (recall that $\alpha = \sqrt{-\lambda}$). All $\lambda \geq 0$ belong to the continuous spectrum.

Let u_k^n (resp. α_k^n) denote the kth eigenfunction (resp. eigenvalue) about the solution W_n. The numerically generated spectra are shown in Table II.

TABLE II

The eigenvalues of the perturbations about the first five solutions W_n obtained numerically. The pseudo-eigenvalue $\alpha = 0$ is also included. The last row corresponding to $n = \infty$ was obtained by solving numerically the transcendental equation (43).

n	α_0^n	α_1^n	α_2^n	α_3^n	α_4^n
0	0				
1	0	4			
2	0	4	27.407		
3	0	4	27.379	182.49	
4	0	4	27.374	182.18	1214.5
...	...	...	...	...	...
∞	0	4	27.37319	182.1202	1210.917

We point out that although $\alpha = 0$ is not a genuine eigenvalue, it is distinguished from the strictly positive part of the continuous spectrum by the fact that the corresponding non-square-integrable pseudo-eigenfunction, called the zero mode, is subdominant at infinity. The existence of the zero mode is due to the time translation symmetry, or in other words, the freedom of shifting the blowup time T in (8). To see this, consider the self-similar solution with a shifted blowup time $W_n((T' - t)/r)$, where $T' = T + \varepsilon$. In terms of the original similarity variables $s = -\ln \sqrt{(T - t)^2 - r^2}$ and $x = (T - t)/r$, we have

$$W_n\left(\frac{T' - t}{r}\right) = W_n(x + \varepsilon e^s \sqrt{x^2 - 1}) = W_n(x) + \varepsilon e^s \sqrt{x^2 - 1}\, W_n'(x) + O(\varepsilon^2)\,,$$

(36)

hence the perturbation generated by shifting the blowup time corresponds to $\alpha = 0$ and has the form

$$u_0^n = \sqrt{x^2 - 1}\, W_n'(x) = \sinh^2\rho\, W_n'(\rho). \tag{37}$$

An alternative way of deriving this result is to take $(\sinh^2\rho\, W')' + 3W(1 - W^2) = 0$, which is (12) reexpressed in terms of ρ, differentiate it and compare with (34).

Since by construction the solution $W_n(\rho)$ has n extrema, it follows from (37) that the zero mode $u_0^n(\rho)$ has n nodes. This implies, by the standard result from Sturm–Liouville theory, that the potential V_n has exactly n negative eigenvalues, in agreement with the numerical results shown in Table II.

We conclude that the self-similar solution W_n has exactly n unstable modes (apart from the unphysical zero mode). In particular, the fundamental solution W_0 is linearly stable, which makes it a candidate for the attractor.

The rest of this section is a digression concerning a striking regularity which an acute reader might have already noticed in Table II. Namely, the third column of Table II indicates that for each $n > 0$ the first eigenvalue below the continuous spectrum $\lambda_1^n = -(\alpha_1^n)^2$ is equal to -16 (with the numerical accuracy of ten decimal places)! This puzzling numerical fact is calling for an explanation. Clearly, it has something to do with the particular form of the nonlinearity since, for instance, the analogous problem for self-similar wave maps from $3 + 1$ dimensional Minkowski spacetime into the 3-sphere does not have this property [4]. We suspect that the problem has some hidden symmetry, yet we cannot exclude a possibility that the numerics is misleading and the eigenvalues λ_1^n are not precisely equal but their splitting is beyond the numerical resolution. Some insight into this puzzle can be gained by analysing the limiting case $n \to \infty$. Recall that $W_n(\rho)$ tends to zero for any $\rho > 0$ as $n \to \infty$, hence the sequence of potentials V_n has the following nonuniform limit

$$\lim_{n \to \infty} V_n(\rho) = V_\infty(\rho) = -\frac{3}{\sinh^2 \rho}. \tag{38}$$

For the limiting potential V_∞ the Schrödinger equation (34) can be solved exactly. The solution that is admissible at infinity (which as before is the limit-point) is given by the associated Legendre function of the first kind

$$u(\rho) = P_\nu^\alpha(\coth \rho), \qquad \nu = -\frac{1}{2} + i\frac{\sqrt{11}}{2}. \tag{39}$$

Here ν is one of the roots of $\nu(\nu + 1) = -3$; the second root gives the same solution because $P_{-1/2+i\beta}^\alpha(x)$ with real β is real and $P_{-1/2+i\beta}^\alpha(x) = P_{-1/2-i\beta}^\alpha(x)$.

As $\rho \to 0$, the solution (39) behaves as

$$u(\rho) \propto \rho^{\frac{1}{2}} \sin\left(\frac{\sqrt{11}}{2} \ln \rho + \delta(\alpha)\right), \tag{40}$$

so it is always admissible, independently of α. This means that $\rho = 0$ is the limit-circle point and, therefore, in order to have a well-defined self-adjoint problem we need to impose an additional boundary condition. In the language of spectral theory such a condition is called a self-adjoint extension. The continuous part of the spectrum is the same for all self-adjoint extensions but the eigenvalues do depend on the choice. In our case the

self-adjoint extension amounts to fixing $\delta(\alpha)$ — the phase of oscillations of the eigenfunctions for $\rho \to 0$. The natural choice is to require that the eigenfunctions oscillate with the same phase as the zero mode, that is $\delta(\alpha) = \delta(0)$, or equivalently

$$\lim_{\rho \to 0} \left\{ P_\nu^0(\coth \rho) u'(\rho) - P_\nu^{0'}(\coth \rho) u(\rho) \right\} = 0. \tag{41}$$

Note that under this condition the following diagram commutes

$$
\begin{array}{ccc}
V_n & \longrightarrow & \{\alpha_k^n\} \\
\downarrow & & \vdots \\
V_\infty & \longrightarrow & \{\alpha_k^\infty\}
\end{array}
$$
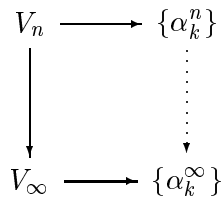
Substituting (39) into (41) and using the asymptotic expansion (for real β)

$$P_{-1/2+i\beta}^\alpha(\coth \rho) \sim \frac{2^{i\beta}\,\Gamma(i\beta)}{\sqrt{2\pi}\,\Gamma(1/2 + i\beta - \alpha)}\,\rho^{1/2+i\beta} + \text{c.c.} \quad \text{for} \quad \rho \to 0, \tag{42}$$

we obtain the quantisation condition for the eigenvalues

$$\arg\left\{ \Gamma\left(\frac{1}{2} - i\frac{\sqrt{11}}{2}\right)\,\Gamma\left(\frac{1}{2} + i\frac{\sqrt{11}}{2} + \alpha_k\right) \right\} = k\pi, \quad k \in \boldsymbol{N}. \tag{43}$$

This transcendental equation has infinitely many roots which for $k \geq 2$ can be obtained only numerically (see the last row in Table II). However, for $k = 1$ the exact solution is $\alpha_1 = 4$ because (accidentally?) $\Gamma(1/2 + i\sqrt{11}/2 + 4) = -45\,\Gamma(1/2 + i\sqrt{11}/2)$, as can be readily verified using four times the identity $\Gamma(z+1) = z\Gamma(z)$. Thus, we showed that the least negative eigenvalue of the limiting potential is equal to -16. Although this analysis does not resolve the original puzzle why all α_1^n are equal to 4, it shows at least that 4 is the accumulation point of this sequence.

The asymptotic distribution of eigenvalues for $k \to \infty$ can be derived from (43) by using the formula for the asymptotic behaviour of the gamma function for large z

$$\Gamma(\nu + z) \sim \sqrt{2\pi}\,e^{(\nu+z-1/2)\ln z - z} \quad \text{for} \quad |z| \to \infty, \tag{44}$$

which yields

$$\arg\left\{ \Gamma\left(\frac{1}{2} + i\frac{\sqrt{11}}{2} + \alpha\right) \right\} \sim \frac{\sqrt{11}}{2}\ln \alpha \quad \text{for} \quad \alpha \to \infty. \tag{45}$$

Applying this to (43) one gets

$$\frac{\alpha_{k+1}}{\alpha_k} \approx e^{\frac{2\pi}{\sqrt{11}}} \quad \text{for } k \to \infty \, . \tag{46}$$

This formula was useful is providing an initial guess in the numerical root finding procedure for equation (43) for large k.

5. Singularities in $D = 5$

Having learned about self-similar solutions, we are now prepared to understand the results of numerical studies, first reported in [7], of the Cauchy problem for the YM equation in five space dimensions

$$w_{tt} = w_{rr} + \frac{2}{r}w_r + \frac{3}{r^2}w(1 - w^2) \, . \tag{47}$$

The main goal of these studies was to determine the asymptotics of blowup. Our numerical simulations were based on finite difference methods combined with adaptive mesh refinement. The latter were instrumental in resolving the structure of singularities developing on vanishingly small scales. We stress that a priori analytical insight into the problem, in particular the knowledge of self-similar solutions was very helpful in interpreting the numerical results.

We solved equation (47) for a variety of initial conditions interpolating between small and large data. A typical example of such initial data is a Gaussian (ingoing or time-symmetric) of the form

$$w(0, r) = 1 - Ar^2 \exp\left[-\sigma(r - R)^2\right] , \tag{48}$$

with adjustable amplitude A and fixed parameters σ and R. The global behaviour of solutions is qualitatively the same for all families of initial data and depends critically on the amplitude A (or any other parameter which controls the "strength" of initial data). For small amplitudes the solutions disperse, that is the energy is radiated away to infinity and in any compact region the solution approaches the vacuum solution $w = 1$. This is in agreement with general theorems on global existence for small initial data [2]. Heuristically, this follows from the fact the for a small amplitude the nonlinearity is dominated by the dispersive effect of the linear wave operator. For large amplitudes we observe the development of two clearly separated regions: an outer region where the evolution is very slow and a rapidly evolving inner region where the solution attains a kink-like shape which shrinks in a self-similar manner to zero size in a finite time T. The kink is, of course, nothing else but the self-similar solution $W_0(\frac{r}{T-t})$[1] (see figure 3). We summarise these findings in the following conjecture:

---

[1] Throughout this section we use the similarity variable $\eta = \frac{r}{T-t}$ (rather than x) and abuse the notation by writing $W_n(\eta)$ to denote $\tilde{W}_n(\eta) = W_n(x)$.

Conjecture 1 (On blowup in $D = 5$). *Solutions of equation (47) corresponding to sufficiently large initial data do blowup in finite time in the sense that $w_{rr}(t, 0)$ diverges as $t \nearrow T$ for some $T > 0$. The universal asymptotic profile of blowup is given by the stable self-similar solution:*

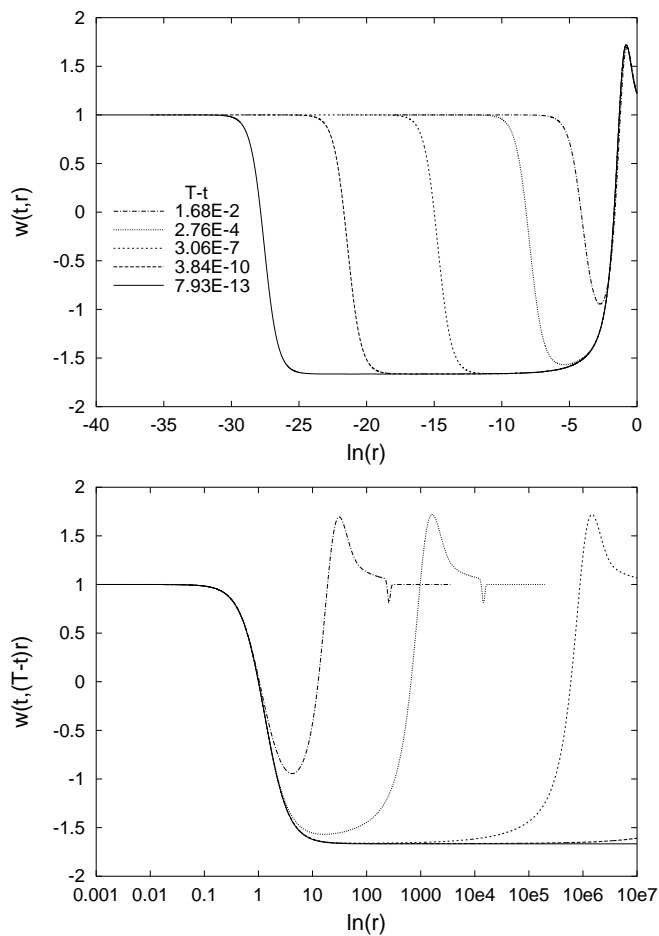$$\lim_{t \nearrow T} w(t, (T - t)r) = W_0(r) \,. \tag{49}$$

Fig. 3. The upper plot shows the late time evolution of time symmetric initial data of the form (48) with $\sigma = 10$, $R = 2$, and $A = 0.2$. As the blowup progresses, the inner solution gradually attains the form of the stable self-similar solution $W_0(r/(T - t))$. The outer solution appears frozen on this timescale. In the lower plot the rescaled solutions $w(t, (T - t))r$ are shown to collapse to the profile $W_0(r)$ (solid line).

We think that the basic mechanism which is responsible for the observed asymptotic self-similarity of blowup can be viewed as the convergence to the lowest "energy" configuration. To see this, let us rewrite (47) in terms of the similarity variable η and the slow time $\tau = -\ln(T - t)$ to get

$$w_{\tau\tau} + w_\tau + 2\eta w_{\eta\tau} = (1 - \eta^2)\left(w_{\eta\eta} + \frac{2}{\eta}w_\eta\right) + \frac{3}{\eta^2}w(1 - w^2). \qquad (50)$$

In this way the problem of blowup was converted into the problem of asymptotic behaviour of solutions for $\tau \to \infty$. The natural "energy" functional associated with this problem is

$$K(w) = \int\limits_0^1 \left(\eta^2 w_\eta^2 + \frac{3}{2}\frac{(1 - w^2)^2}{1 - \eta^2}\right) d\eta. \qquad (51)$$

$K(w)$ has a minimum at the self-similar solution W_0 and saddle points with n unstable directions at solutions W_n with $n > 0$. Since the wave equation (50) contains a damping term reflecting an outward flux of energy through the past light cone of the singularity, we suspect (but cannot prove) that $K(w)$ decreases with time. If so, it is natural to expect that solutions will tend asymptotically to the minimum of $K(w)$.

We already know that solutions with small data disperse and solutions with large data blow up. The question is what happens in between. Using bisection, we found that along each interpolating family of initial data there is a threshold value of the parameter, say the amplitude A^*, below which the solutions disperse and above which a singularity is formed. The evolution of initial data near the threshold was found to go through a transient phase which is universal, *i.e.* the same for all families. This intermediate attractor was identified as the self-similar solution W_1. Having gone through this transient phase, at the end the solutions leave the intermediate attractor towards dispersal or blowup. This behaviour is shown in figure 4 for the time-symmetric initial data of the form (48).

The universality of the dynamics at the threshold of singularity formation can be understood heuristically as follows[2]. As we showed above, the self-similar solution W_1 has exactly one unstable mode — in other words the stable manifold of this solution has codimension one and therefore generic one-parameter families of initial data do intersect it. The points of intersection correspond to critical initial data that converge asymptotically to W_1. The marginally critical data, by continuity, initially remain close to the

---

[2] This heuristic picture of the dynamics near the threshold, borrowed from dynamical systems theory, has been first given in the context of Einstein's equation — see Section 6 and [12].
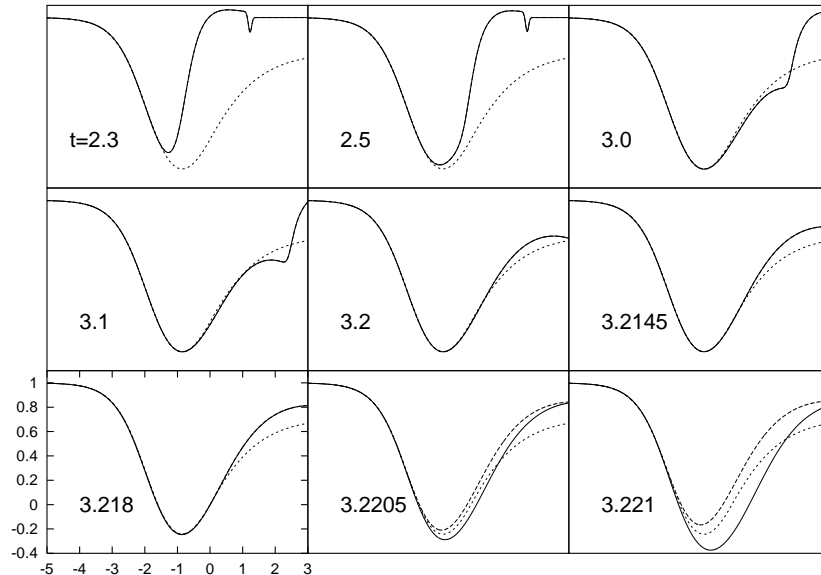
Fig. 4. The dynamics of time-symmetric initial data of the form (48) with ampli-tudes that are fine-tuned to the threshold of singularity formation. The rescaled solution $w(t, (T - t)r)$ is plotted against $\ln(r)$ for a sequence of intermediate times. Shown (solid and dashed lines) is the pair of solutions starting with marginally crit-ical amplitudes $A = A^* \pm \varepsilon$, where $A^* = 0.144296087005405$. Since $\varepsilon = 10^{-15}$, the two solutions are indistinguishable on the first seven frames. The convergence to the self-similar solution W_1 (dotted line) is clearly seen in the intermediate asymp-totics. The last two frames show the solutions departing from the intermediate attractor towards blowup and dispersal, respectively.

stable manifold and approach W_1 for intermediate times but eventually are repelled from its vicinity along the one-dimensional unstable manifold (see figure 5). According to this picture the universality of the nearly critical dynamics follows immediately from the fact that the same unstable mode dominates the evolution of all solutions. More precisely, the evolution of marginally critical solutions in the intermediate asymptotics can be approx-imated as

$$w(t, r) = W_1(\eta) + c(A)(T - t)^{-\alpha_1 - 1}v_(\eta) + \text{radiation}, \qquad (52)$$

where v_1 is the single unstable mode with the eigenvalue $\alpha_1 = 4$. The small constant $c(A)$, which is the only vestige of the initial data, quantifies an admixture of the unstable mode — for precisely critical data $c(A^*) = 0$. The time of departure from the intermediate attractor is determined by the time t^* in which the unstable mode grows to a finite size, $i.e.$,
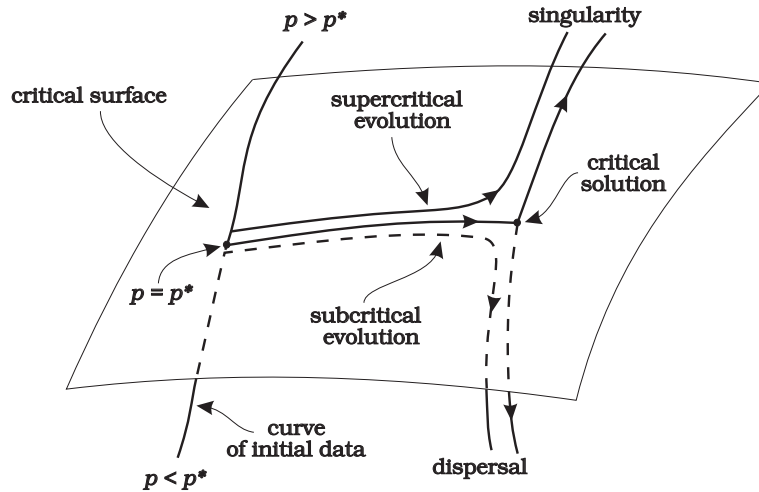
Fig. 5. A schematic phase space picture of the dynamics at the threshold of singularity formation.

$c(A)(T - t^*)^{-\alpha_1 - 1} \sim O(1)$. Using $c(A) \approx c'(A^*)(A - A^*)$ and substituting $\alpha_1 = 4$, we get $T - t^* \sim |A^* - A|^{1/5}$. Various scaling laws can be derived from this. For example, consider solutions with marginally sub-threshold amplitudes $A = A^* - \varepsilon$. For such solutions the energy density

$$e(t, r) = \frac{w_t^2}{r^2} + \frac{w_r^2}{r^2} + \frac{3(1 - w^2)^2}{2r^4} \tag{53}$$

initially grows at the centre, attains a maximum at a certain time $\approx t^*$ and then drops to zero. Substituting (52) into (53) we get that $e(t, 0) \sim (T - t)^{-4}$, and hence $e(t^*, 0) \sim \varepsilon^{-4/5}$.

6. Connection with critical phenomena in gravitational collapse

The behaviour of solutions near the threshold of singularity formation described above shares many features with critical phenomena at the threshold of black hole formation in gravitational collapse. To explain these similarities, we now briefly recall the phenomenology and heuristics of the critical gravitational collapse. Consider a spherical shell of matter and let it collapse under its own weight. The dynamics of this process, modelled by Einstein's equations, can be understood intuitively in terms of the competition between gravitational attraction and repulsive internal forces (due, for instance, to kinetic energy of matter or pressure). If the initial configuration is dilute, then the repulsive forces "win" and the collapsing matter will rebound or implode through the centre, and eventually will disperse. On

the other hand, if the density of matter is sufficiently large, some fraction of the initial mass will form a black hole. Critical gravitational collapse occurs when the attracting and repulsive forces governing the dynamics of this process are almost in balance, or in other words, the initial configuration is near the threshold of black hole formation. The systematic studies of critical gravitational collapse were launched in the early nineties by the seminal paper by Choptuik [11] in which he investigated numerically the collapse of a self-gravitating massless scalar field.

Evolving initial data fine-tuned to the border between no-black-hole and black-hole spacetimes, Choptuik found the following unforseen phenomena near the threshold:

(i) universality: all initial data which are near the black hole threshold go through a universal transient period in their evolution during which they approach a certain intermediate attractor, before eventually dispersing or forming a black hole. This universal intermediate attractor is usually referred to as the critical solution.

(ii) discrete self-similarity: the critical solution is discretely self-similar, that is it is invariant under dilations by a certain fixed factor called the echoing period.

(iii) black-hole mass scaling: for initial data that do form black holes, the masses of black holes satisfy the power law $M_{\mathrm{bh}} \sim \varepsilon^{\gamma}$, where ε is the distance to the threshold and γ is a universal (i.e., the same for all initial data) critical exponent. Thus, by fine tuning to the threshold one can make an arbitrarily tiny black hole. Put differently, there is no mass gap at the transition between black-hole and no-black-hole spacetimes.

What Choptuik found for the scalar field, has been later observed in many other models of gravitational collapse, although the symmetry of the critical solution itself was found to depend on the model: in some cases the critical solution is self-similar (continuously or discretely), while in other cases the critical solution is static (or periodic). In the latter case black hole formation turns on with finite mass. These two kinds of critical behaviour are referred to as the type II or type I criticality, respectively, to emphasise the formal analogy with second and first order phase transitions in statistical physics. We refer the interested reader to [12] for an excellent review of the growing literature on critical gravitational collapse.

The present understanding of critical behaviour in gravitational collapse is based on the same phase space picture as in figure 5, that is, it is associated with the existence of a critical solution with exactly one unstable mode. This picture leads to some quantitative predictions. In particular, in the case of

type II critical collapse, an elementary dimensional analysis shows that the critical exponent γ in the power law $M_{\mathrm{bh}} \sim \varepsilon^{\gamma}$ is a reciprocal of the unstable eigenvalue of the critical solution.

By now, the similarities between type II critical gravitational collapse and the dynamics at the threshold of singularity formation in the $5 + 1$ YM equations should be evident. This analogy, together with similar results for wave maps in $3 + 1$ dimensions [5,13], shows that the basic properties of critical collapse, such as universality, scaling, and self-similarity, first observed for Einstein's equations, actually have nothing to do with gravity and seem to be robust properties of supercritical nonlinear wave equations. The obvious advantage of toy models, such as the one presented in this paper, is their simplicity which allowed to get a much better analytic grip on critical phenomena than in the case of Einstein's equations; in particular, it was possible to prove existence of the critical solution. The only characteristic property of type II critical collapse which so far has not found in simpler models (besides, of course, the absence of black holes which are replaced by singularities) is discrete self-similarity of the critical solution. It would be very interesting to design a toy model which exhibits discrete self-similarity at the threshold for singularity formation because this could give us insight into the origin of this mysterious symmetry.

7. Singularities in $D = 4$

In this section we consider the Cauchy problem for the YM equation in four space dimensions

$$w_{tt} = w_{rr} + \frac{1}{r}w_r + \frac{2}{r^2}w(1 - w^2). \tag{54}$$

We begin by recalling some facts concerning equation (54) which are be important in understanding the dynamics of singularity formation. First, we note that, in contrast to $D = 5$, there are no smooth self-similar solution in $D = 4$. This follows from the fact that in D dimensions the local solutions of equation (10) near the past light cone behave as $(1 - \eta^2)^{\frac{D-3}{2}}$, hence they are not smooth if D is even (in particular, they are not differentiable in $D = 4$). Although such singular self-similar solutions do exist, they cannot develop from smooth initial data and therefore they are not expected to participate in the dynamics.

Second, $D = 4$ is the critical dimension in the sense that the energy (7) does not change under scaling. This means that, even though the model is scale invariant, a nontrivial finite energy static solution may exist[3]. In fact,

---

[3] Another way of seeing this is to notice that only in $D = 4$ the YM coupling constant e^2 provides the scale of energy.

such a static solution is well known

$$W_{\mathrm{S}}(r) = \frac{1 - r^2}{1 + r^2} \,. \tag{55}$$

This is the instanton in the four-dimensional Euclidean YM theory. Of course, by reflection symmetry, $-W_{\mathrm{S}}(r)$ is also the solution. Since the model is scale invariant, the solution $W_{\mathrm{S}}(r)$ generates an orbit of static solutions $W_{\mathrm{S}}^{\lambda}(r) = W_{\mathrm{S}}(r/\lambda)$, where $0 < \lambda < \infty$.

To analyse the linear stability of the instanton, we insert $w(t,r) = W_{\mathrm{S}}(r) + \mathrm{e}^{ikt}v(r)$ into (54) and linearise. In this way we get the eigenvalue problem (the radial Schrödinger equation)

$$\left(-\frac{d^2}{dr^2} - \frac{1}{r}\frac{d}{dr} + V(r) \right) v = k^2 v, \qquad V(r) = -\frac{2(1 - 3W_{\mathrm{S}}^2)}{r^2} \,. \tag{56}$$

This problem has a zero eigenvalue $k^2 = 0$ which follows from scale invariance. The corresponding eigenfunction (so called zero mode) is determined by the perturbation generated by scaling

$$v_0(r) = -\frac{d}{d\lambda}W_{\mathrm{S}}^{\lambda}(r)\Big|_{\lambda=1} = rW_{\mathrm{S}}'(r) = \frac{4r^2}{(1 + r^2)^2} \,. \tag{57}$$

Since the zero mode $v_0(r)$ has no nodes, it follows by the standard result from Sturm–Liouville theory that there are no negative eigenvalues, and *eo ipso* no unstable modes around $W_{\mathrm{S}}(r)$. Thus, the instanton is marginally stable. Note that the zero eigenvalue lies at the bottom of the continuous spectrum $k^2 \geq 0$, hence there is no spectral gap in the problem.

After these preliminaries, we return to the discussion of the Cauchy problem for equation (54). For small energies the solutions disperse, in agreement with general theorems. For large energies, at first sight the global behaviour seems similar to the $D = 5$ case — as before, near the centre the solution attains the form of a kink which shrinks to zero size. However, this similarity is superficial because now the kink is not a self-similar solution (as no such solution exists). It turns out (see figure 6) that the kink has the form of the *scale-evolving* instanton

$$w(t,r) \approx W_{\mathrm{S}}\left(\frac{r}{\lambda(t)} \right), \tag{58}$$

where a scaling factor $\lambda(t)$ is a positive function which tends to zero as $t \to T$. We summarise these findings in the following conjecture:

Conjecture 2 (On blowup in $D = 4$). *Solutions of equation (54) with sufficiently large energy do blow up in finite time in the sense that $w_{rr}(t, 0)$ diverges as $t \nearrow T$ for some $T > 0$. The universal asymptotic profile of blowup is given by the instanton. More precisely, there exists a positive function $\lambda(t) \searrow 0$ for $t \nearrow T$ such that*

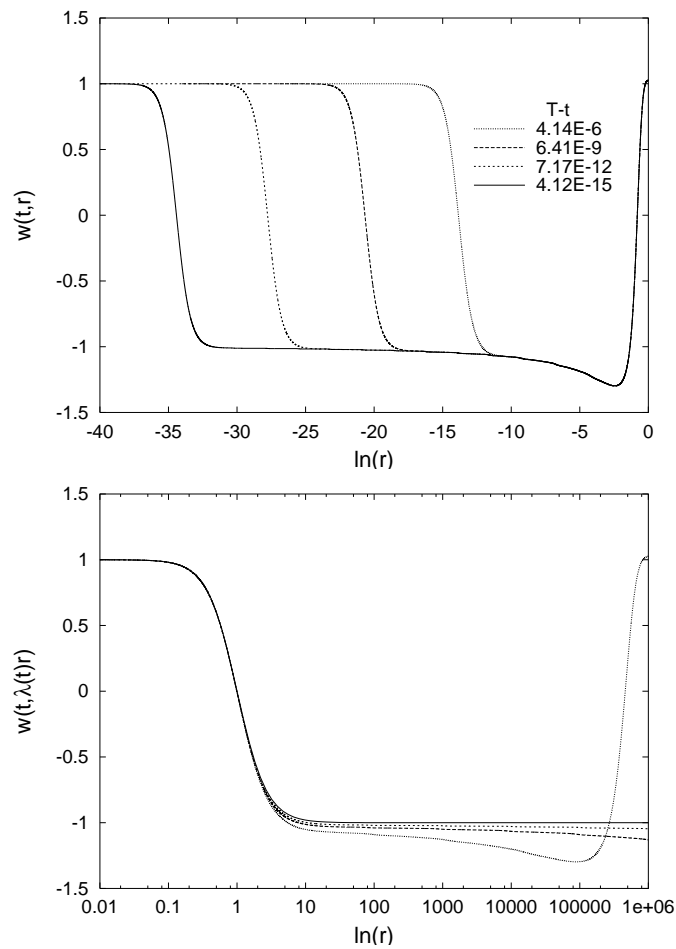$$\lim_{t \nearrow T} u(t, \lambda(t)r) = W_{\mathrm{S}}(r). \qquad (59)$$

Fig. 6. The upper plot shows the formation of a singularity for large initial data of the form (48) with $A = 0.5$. The inner solution has the form of the scale-evolving instanton $W_{\mathrm{S}}(r/\lambda(t))$ with the scale factor $\lambda(t)$ going to zero slightly faster than linearly. In the lower plot the rescaled solutions are shown to collapse to the profile of the instanton $W_{\mathrm{S}}(r)$ (solid line).

The key question which is left open in this conjecture is: what determines the evolution of the scaling factor $\lambda(t)$; in particular, what is the asymptotic behaviour of $\lambda(t)$ for $t \to T$? Numerical evidence shown in figure 7 suggests that the rate of blowup goes asymptotically to zero, that is ($\dot{} = d/dt$)

$$\lim_{t \to T} \frac{\lambda(t)}{T - t} = -\lim_{t \to T} \dot{\lambda} = 0, \tag{60}$$

but it seems very hard to determine an exact asymptotics of $\lambda(t)$ from pure numerics[4].
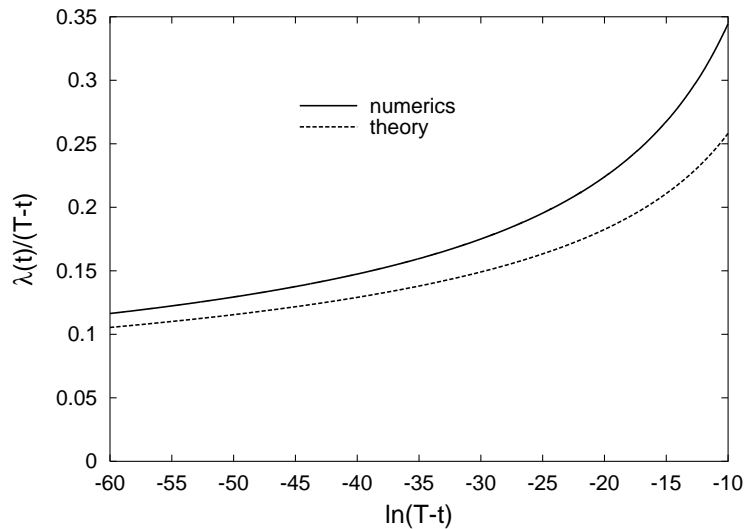

Fig. 7. Comparison of the numerically computed scaling factor divided by $T - t$ (for the same data as in figure 6) with the analytic formula $\frac{\lambda(t)}{T-t} = \sqrt{\frac{2}{3}} (-\ln(T-t))^{-1/2}$.

Recently, an analytical approach to this problem has been suggested in [15]. Below we sketch the main idea of this approach. Let $\mathcal{M} \equiv \{W_{\mathrm{S}}(r/\lambda) | \lambda \in \mathbb{R}^+\}$ be a manifold of rescaled instantons (a one-dimensional centre manifold). Assuming that a solution is in a neighbourhood of $\mathcal{M}$, we decompose it as

$$w = W_{\mathrm{S}}(\eta) + v(t, \eta), \qquad \eta = r/\lambda(t). \tag{61}$$

Here v represents a small deviation of the solution from $\mathcal{M}$ and λ is the collective coordinate on $\mathcal{M}$. To fix the splitting between these two parts we

---

[4] This issue is also discussed by Linhart and Sadun in [14].

require that $Pv = 0$, where P is the projection on $\mathcal{M}$. Plugging (61) into (54) we get ($' = d/d\eta$)

$$\lambda^2 \ddot{v} - 2\eta \dot{\lambda} \lambda \dot{v}' + Lv + N(v) = \lambda \ddot{\lambda} \eta W'_{\mathrm{S}} - \dot{\lambda}^2 (\eta^2 W''_{\mathrm{S}} + 2\eta W'_{\mathrm{S}}), \qquad (62)$$

where L is the linear perturbation operator about the instanton

$$L = -\frac{\partial^2}{\partial \eta^2} - \frac{1}{\eta}\frac{\partial}{\partial \eta} - \frac{2(1 - 3W_{\mathrm{S}}^2)}{\eta^2}, \qquad (63)$$

and

$$N(v) = \frac{6W_{\mathrm{S}}}{\eta^2} v^2 + \frac{2}{\eta^2} v^3. \qquad (64)$$

It is clear from (62) that $v = O(\dot{\lambda}^2)$, hence for v to decay to zero as $t \to T$, the rate of blowup must go to zero as well. We stress this point to emphasise that the linear evolution of $\lambda(t)$, predicted for example by the geodesic approximation, is inconsistent with Conjecture 2. Next, by projecting equation (62) on $\mathcal{M}$ and in the orthogonal direction, we get a coupled system consisting of a nonhomogeneous wave equation for v and an ordinary differential equation for λ. Solving the first equation for v and plugging the result into the second equation, we obtain in the lowest order the following modulation equation

$$\lambda \ddot{\lambda} = \frac{3}{4} \dot{\lambda}^4. \qquad (65)$$

From this we get the leading order asymptotics for $t \to T$

$$\lambda(t) \sim \sqrt{\frac{2}{3}} \frac{T - t}{\sqrt{-\ln(T - t)}}. \qquad (66)$$

As shown in figure 7 this result is in rough agreement with numerics. There are many possible sources of the apparent discrepancy. On the numerical side there are discretisation errors, an error in estimating the blowup time, or errors in computing λ from the data. On the analytical side, there might be corrections to (66) coming from the bounded region expansion and, more importantly, from the far field behaviour[5]. Finally, and in our opinion most likely reason of discrepancy is that the solution shown in figure 7 has not yet reached the truly asymptotic regime and consequently the higher order corrections to formula (66) are still significant.

---

[5] The derivation of (65) is not quite straightforward because of the presence of infrared divergencies which need to be regularised.

The issue of blowup rate is closely related to the problem of energy concentration in the singularity. To explain this, we define the kinetic and the potential energies at time $t < T$ inside the past light cone of the singularity

$$E_{\mathrm{K}}(t) = 6\pi^2 \int\limits_0^{T-t} w_t^2 \, r \, dr, \qquad E_{\mathrm{P}}(t) = 6\pi^2 \int\limits_0^{T-t} \left(w_r^2 + \frac{(1-w^2)^2}{r^2} \right) r \, dr. \quad (67)$$

Substituting (58) into (67) we obtain

$$E_{\mathrm{K}}(t) = 6\pi^2 \dot\lambda^2 \int\limits_0^{\frac{T-t}{\lambda(t)}} W_{\mathrm{S}}'^2 \, r \, dr, \qquad E_{\mathrm{P}}(t) = 6\pi^2 \int\limits_0^{\frac{T-t}{\lambda(t)}} \left(W_{\mathrm{S}}'^2 + \frac{(1-W_{\mathrm{S}}^2)^2}{r^2} \right) r \, dr.$$
$$(68)$$

Assuming (60), this implies that

$$\lim_{t\to T} E_{\mathrm{K}}(t) = 0, \qquad \lim_{t\to T} E_{\mathrm{P}}(t) = 16\pi^2. \qquad (69)$$

Thus, the energy equal to the energy of the instanton gets concentrated in the singularity. This means that in the process of blowup the excess energy must be radiated away from the inner region as the solution converges to the instanton.

It is worth pointing out that the concentration of energy is a necessary condition for blowup in the critical dimension. To see this, suppose that the solution blows up at time T and assume for contradiction that $\lim_{t\to T} E(t) = 0$. Then, by choosing a sufficiently small $\varepsilon > 0$ we can have $E(T - \varepsilon)$ arbitrarily small. This implies, by causality and global existence for small energy data, that the solution exists globally in time, contradicting the assumption. For radial equations of the type

$$w_{tt} = w_{rr} + \frac{1}{r} w_r + \frac{f(w)}{r^2}, \qquad (70)$$

a stronger result was proved by Struwe [16] who basically showed that if the solution of (70) blows up, then it must do so in the manner described in Conjecture 2. In this sense formation of singularities is intimately tied with the existence of a static solution.

Finally, we address briefly the issue of the threshold of singularity formation. Using the technique described in Section 5, along each interpolating family of initial data we can determine the critical point separating blowup from dispersal. However, in contrast to the $D = 5$ case, we see no evidence for the existence of an intermediate attractor in the evolution of nearly critical data. This fact, together with a similar result for $2 + 1$ dimensional

wave maps [6] suggests that in the critical dimension the transition between blowup and dispersal is not governed by any critical solution. We suspect that the evolution of precisely critical initial data still has the form (58) but the dynamics of the scaling factor is different than in (66). This belief is based on the fact that in the evolution of marginally critical initial data we can clearly distinguish a transient phase during which $\lambda(t)$ drops very quickly and then, either reaches a minimum and starts growing (in the case of dispersal), or keeps decreasing with the "normal" rate (in the case of blowup).

8. Conclusions

There are two main lessons that we wanted to convey in this survey. The first lesson is that there are striking analogies between major evolution equations. In particular, the mechanism of blowup to a large extent is determined by the criticality class of the model. These analogies can be used to get insight into hard problems (such as singularity formation for Einstein's equations) by studying toy models which belong to the *same* criticality class. This approach is in the spirit of general philosophy expressed by David Hilbert in his famous lecture delivered before the International Congress of Mathematicians at Paris in 1900 [17]: *"In dealing with mathematical problems, specialisation plays, as I believe, a still more important part than generalisation. Perhaps in most cases where we seek in vain the answer to a question, the cause of the failure lies in the fact that problems simpler and easier than the one in hand have been either not at all or incompletely solved. All depends, then, on finding out these easier problems, and on solving them by means of devices as perfect as possible and of concepts capable of generalisation."*

The second lesson is concerned with the interplay between numerical and analytical techniques. Accurate and reliable numerical simulation of singular behaviour is difficult and hard to assess. In order to keep track of a singularity developing on exceedingly small spatio-temporal scales, one needs sophisticated techniques such as adaptive mesh refinement. For these techniques the convergence and error analysis are lacking so extreme care is needed to make sure that the computed singularities are not numerical artifacts. For this reason, in order to feel confident about numerics it is important to have some analytical information, like existence of self-similar solutions. Without a theory, simulations alone do not provide ample evidence for the existence of a singularity. We believe that the interaction between numerical and analytical techniques, illustrated here by the studies of blowup, will become more and more important in future as we begin to attack more difficult problems.

I am grateful to Michael Sigal and Yu. N. Ovchinnikov for permission to announce here the results of yet unpublished joint work. I thank Zbisław Tabor for providing me with data for figure 7. I acknowledge the hospitality of the Albert Einstein Institute for Gravitational Physics in Golm and the Erwin Schrödinger Institute for Mathematical Physics in Vienna, where parts of this paper were produced.

REFERENCES

[1] D. Eardley, V. Moncrief, *Commun. Math. Phys.* **83**, 171 (1982).

[2] S. Klainerman, in: *Progress in Nonlinear Differential Equations and Their Applications*, Birkhäuser, 1997, vol. 29.

[3] T. Cazenave, J. Shatah, A. Shadi Tahvildar-Zadeh, *Ann. Inst. Henri Poincare* **68**, 315 (1998).

[4] P. Bizoń, *Commun. Math. Phys.* **215**, 45 (2000).

[5] P. Bizoń, T. Chmaj, Z. Tabor, *Nonlinearity* **13**, 1411 (2000).

[6] P. Bizoń, T. Chmaj, Z. Tabor, *Nonlinearity* **14**, 1041 (2001).

[7] P. Bizoń, Z. Tabor, *Phys. Rev.* **D64**, 121701 (2001).

[8] O. Dumitrascu, *Stud. Cercet. Mat.* **34**(4), 329 (1982).

[9] S. Klainerman, M. Machedon, *Ann. Math.* **142**, 39 (1995).

[10] P. Breitenlohner, P. Forgács, D. Maison, *Commun. Math. Phys.* **163**, 141 (1994).

[11] M.W. Choptuik, *Phys. Rev. Lett.* **70**, 9 (1993).

[12] C. Gundlach, *Living Rev. Relativity* **2**, (4) (1999).

[13] S.L. Liebling, E. Hirschmann, J. Isenberg, *J. Math. Phys.* **41**, 5691 (2000).

[14] J.M. Linhart, L.A. Sadun, *Nonlinearity* **15**, 219 (2002).

[15] P. Bizoń, Yu.N. Ovchinnikov, I.M. Sigal, in preparation.

[16] M. Struwe, preprint (2000), to appear in *Commun. Pure Appl. Math.*

[17] D. Hilbert, reprinted in *Bull. Amer. Math. Soc.* **37**, 407 (2000).