MEMOIRS
of the
American Mathematical Society

Volume 221 • Number 1041 (fifth of 5 numbers) • January 2013

The Regularity of General Parabolic Systems with Degenerate Diffusion

Verena Bögelein
Frank Duzaar
Giuseppe Mingione

ISSN 0065-9266 (print) ISSN 1947-6221 (online)

American Mathematical Society

MEMOIRS
of the
American Mathematical Society

Number 1041

The Regularity of General Parabolic Systems with Degenerate Diffusion

Verena Bögelein
Frank Duzaar
Giuseppe Mingione

January 2013 • Volume 221 • Number 1041 (fifth of 5 numbers) • ISSN 0065-9266

American Mathematical Society
Providence, Rhode Island

Library of Congress Cataloging-in-Publication Data

Bögelein, Verena, 1978-

The regularity of general parabolic systems with degenerate diffusion / Verena Bögelein, Frank Duzaar, Giuseppe Mingione.

p. cm. – (Memoirs of the American Mathematical Society, ISSN 0065-9266 ; number 1041)

"January 2013, volume 221, number 1041 (end of volume)."

Includes bibliographical references.

ISBN 978-0-8218-8975-6 (alk. paper)

1. Degenerate differential equations. 2. Differential equations, Parabolic. 3. Diffusion–Mathematical models. I. Duzaar, Frank, 1957- II Mingione, Giuseppe, 1972- III. Title.

QA377.5.B64 2013

515′.3534–dc23 2012035243

Memoirs of the American Mathematical Society

This journal is devoted entirely to research in pure and applied mathematics.

Publisher Item Identifier. The Publisher Item Identifier (PII) appears as a footnote on the Abstract page of each article. This alphanumeric string of characters uniquely identifies each article and can be used for future cataloguing, searching, and electronic retrieval.

Subscription information. Beginning with the January 2010 issue, *Memoirs* is accessible from `www.ams.org/journals`. The 2013 subscription begins with volume 221 and consists of six mailings, each containing one or more numbers. Subscription prices are as follows: for paper delivery, US$795 list, US$636 institutional member; for electronic delivery, US$700 list, US$560 institutional member. Upon request, subscribers to paper delivery of this journal are also entitled to receive electronic delivery. If ordering the paper version, add US$10 for delivery within the United States; US$69 for outside the United States. Subscription renewals are subject to late fees. See `www.ams.org/help-faq` for more journal subscription information. Each number may be ordered separately; *please specify number* when ordering an individual number.

Back number information. For back issues see `www.ams.org/bookstore`.

Subscriptions and orders should be addressed to the American Mathematical Society, P. O. Box 845904, Boston, MA 02284-5904 USA. *All orders must be accompanied by payment.* Other correspondence should be addressed to 201 Charles Street, Providence, RI 02904-2294 USA.

Memoirs of the American Mathematical Society (ISSN 0065-9266) is published bimonthly (each volume consisting usually of more than one number) by the American Mathematical Society at 201 Charles Street, Providence, RI 02904-2294 USA. Periodicals postage paid at Providence, RI. Postmaster: Send address changes to Memoirs, American Mathematical Society, 201 Charles Street, Providence, RI 02904-2294 USA.

This publication is indexed in *Mathematical Reviews*®, *Zentralblatt MATH*, *Science Citation Index*®, *Science Citation Index*™*-Expanded*, *ISI Alerting Services*℠, *SciSearch*®, *Research Alert*®, *CompuMath Citation Index*®, *Current Contents*®*/Physical, Chemical & Earth Sciences*. This publication is archived in
Portico and *CLOCKSS*.
Printed in the United States of America.

♾ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1 18 17 16 15 14 13

Abstract

The aim of the paper is twofold. On one hand we want to present a new technique called p-caloric approximation, which is a proper generalization of the classical compactness methods first developed by DeGiorgi with his Harmonic Approximation Lemma. This last result, initially introduced in the setting of Geometric Measure Theory to prove the regularity of minimal surfaces, is nowadays a classical tool to prove linearization and regularity results for vectorial problems. Here we develop a very far reaching version of this general principle devised to linearize general degenerate parabolic systems. The use of this result in turn allows to achieve the subsequent and main aim of the paper, that is the implementation of a partial regularity theory for parabolic systems with degenerate diffusion of the type

$$\partial_t u - \operatorname{div} a(Du) = 0\,, \tag{0.1}$$

without necessarily assuming a quasi-diagonal structure, i.e. a structure prescribing that the gradient non-linearities depend only on the explicit scalar quantity $|Du|$. Indeed, the by now classical theory of DiBenedetto (*Degenerate parabolic equations*, Universitext, New York, NY, Springer-Verlag, 1993) introduces the fundamental concept of intrinsic geometry and allows to deal with the classical degenerate parabolic p-Laplacian system

$$\partial_t u - \operatorname{div}(|Du|^{p-2}Du) = 0 \tag{0.2}$$

and more generally with systems of the type

$$\partial_t u - \operatorname{div}\big(g(|Du|)Du\big) = 0\,. \tag{0.3}$$

Here, we take such regularity results as a starting point and develop a partial regularity theory – regularity of solutions outside a negligible closed subset of the domain – applying to general degenerate parabolic systems of the type (0.1), thereby

Received by the editor April 1, 2010.

Article electronically published on May 15, 2012; S 0065-9266(2012)00664-2.

2010 *Mathematics Subject Classification.* 35K55, 35K65, 35K67.

Key words and phrases. Partial regularity, degenerate parabolic systems, singular parabolic systems, p-caloric approximation.

This work is supported by the ERC grant 207573 "Vectorial Problems".

Affiliations at time of publication: V. Bögelein, Department Mathematik, Universität Erlangen–Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany, email: boegelein@mi.uni-erlangen.de; F. Duzaar, Department Mathematik, Universität Erlangen–Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany, email: duzaar@mi.uni-erlangen.de; and G. Mingione, Dipartimento di Matematica, Università di Parma, Viale Usberti 53/a, Campus, 43100 Parma, Italy, email: giuseppe.mingione@unipr.it.

not necessarily satisfying a structure assumption as (0.3). The partial regularity rather than the everywhere one, is natural since even in the non-degenerate case, when considering systems with general structure, singularities may occur. The proof of the almost everywhere regularity of solutions is then achieved via an extremely delicate combination of local linearization methods, together with a proper use of DiBenedetto's intrinsic geometry: the general approach that consists in performing the local analysis by considering parabolic cylinders whose space-time scaling depend on the local behavior of the solution itself. The combination of these approaches was exactly the missing link to prove partial regularity for general parabolic systems considered in (0.1). In turn, the implementation realizing such a matching between the two existing theories is made possible by the p-caloric approximation lemma. More precisely, the proof involves two different kinds of linearization techniques: a more traditional one in those zones where the system is non-degenerate and the original solution is locally compared to solutions of a suitable linear system, and a degenerate one in the zones where the system is truly degenerate and the solution can be compared with solutions of systems as (0.2) via the p-caloric approximation lemma.

Contents

CHAPTER 1

Introduction and Results

The aim of this paper is twofold: the first, and main one, is to establish a rather satisfying regularity theory for general parabolic systems with degenerate diffusion; the second one, of more technical and specialized character, is to introduce a suitable analog of the classical harmonic approximation methods of DeGiorgi's pioneering work [**14**] from the elliptic setting, which in turn is the key to a regularity theory for degenerate parabolic systems.

1.1. A short introduction to the regularity of parabolic systems

A brief description of the present status of the regularity theory for general parabolic systems cannot begin but with the observation that already in the case of elliptic systems the so called *partial regularity* – also called almost everywhere regularity – is in general the best one can usually expect, and therefore the same happens in the case of parabolic systems. Indeed, since the important counterexample of DeGiorgi [**15**] – see also [**50, 51, 59**] – it is known that when dealing with general elliptic systems of the type

$$\operatorname{div} a(Du) = 0 \qquad \text{or} \qquad \operatorname{div}(A(x)Du) = 0$$

considered in an open subset $\Omega \subseteq \mathbb{R}^n$, solutions might possess singularities, and therefore everywhere regularity fails to hold in general. Instead, one can show partial regularity of solutions, i.e. they are regular outside a negligible closed subset, thereby called the singular set of the solution:

$$u \in C^{1,\alpha}_{\mathrm{loc}}(\Omega_u, \mathbb{R}^N) \qquad \text{and} \qquad |\Omega \setminus \Omega_u| = 0 \tag{1.1}$$

and we refer to [**33, 35, 48**] for an account of the theory and a list of references. Eventually estimates for the Hausdorff dimension and boundary regularity can be inferred [**46, 47, 24**]. Let us mention that related results for integral functionals in the calculus of variations are obtained in [**40, 41**]. The above partial regularity results for elliptic systems have been extended to the case of parabolic systems of the type

$$\partial_t u = \operatorname{div} a(Du) \tag{1.2}$$

and we refer for instance to [**6, 8, 27, 29, 53, 54**] for the most recent and sharp theorems on the issue. Let us meanwhile remark that the system in (1.2), as all the other parabolic ones in this paper, will be considered in the cylindrical domain

$$\Omega_T := \Omega \times (0, T)\,,$$

where $\Omega \subseteq \mathbb{R}^n$ is an open bounded domain with $n \geq 2$. Such partial regularity results are however obtained under a non-degenerate ellipticity assumption, both in the case of systems and in that of variational integrals, and this amounts to require – when considering problems with so called p-growth, that is when we prescribe on $a(Du)$ a growth bound in terms of $|Du|^{p-1}$ – that

$$\nu(1+|q|^2)^{\frac{p-2}{2}}|\xi|^2 \leq \langle Da(q)\xi,\xi\rangle \qquad \text{for every } q,\xi \in \mathbb{R}^{Nn}\,. \tag{1.3}$$

We are of course assuming that u takes its values in $\mathbb{R}^N$ and $N > 1$. An assumption like (1.3) is however not satisfied for important examples as for instance the p-Laplacian system

$$\operatorname{div}(|Du|^{p-2}Du) = 0 \tag{1.4}$$

and its evolutionary version, the parabolic p-Laplacian system:

$$\partial_t u = \operatorname{div}(|Du|^{p-2}Du)\,. \tag{1.5}$$

The first important observation is that, both the system in (1.4) and the one in (1.5), present a central, additional feature: the gradient non-linearity only depends on the scalar quantity $|Du|$. As a matter of fact the interior regularity results available for (1.5) readily extend to systems of the type

$$\partial_t u = \operatorname{div}(g(|Du|)Du)\,, \tag{1.6}$$

assuming that

$$g(|Du|) \approx |Du|^{p-2}\,, \tag{1.7}$$

where the symbol $\approx$ must be made precise in a suitable way. Assumption (1.7) tells, once again, that the gradient non-linearity depends on Du via the modulus $|Du|$. This is in fact of crucial importance, already in the elliptic case, in order to pass from partial regularity to everywhere regularity and proving that the singular set is in fact empty. Indeed, an approach to the regularity of solutions to (1.4) has been given for the first time by Uhlenbeck [**60**] - the scalar case $N = 1$ had been previously treated by Ural'tseva [**61**] - who proved that solutions to (1.4) are locally of class $C^{1,\alpha}$ for some $\alpha \in (0,1)$. We remark that up to now the structure condition

$$a(Du) = g(|Du|)Du$$

is essentially the only one known to imply the everywhere interior regularity of solutions, and this fact goes back to the work of Uhlenbeck, as just mentioned; for a recent updated survey we refer for instance to [**48**].

The regularity theory for evolutionary systems of the type (1.5) *and* (1.6) *is instead a fundamental achievement of DiBenedetto & Friedman*, which is reported in the papers [**18, 19, 20**], where the concept of intrinsic geometry has been extensively used in order to obtain the relevant local estimates. For the $C^{1,\alpha}$-estimate see also [**62**]. The intrinsic geometry approach of DiBenedetto [**16**], which is by now classical and that is described at length in the monograph [**17**], is actually at the origin of virtually all the techniques developed up to now to prove regularity results for degenerate parabolic problems, see for instance [**3, 5, 7, 21, 38, 39, 52**]. This approach prescribes, roughly speaking, that the regularity of solutions to evolutionary equations as in (1.5) has to be studied by considering the behavior of solutions on shrinking cylinders whose space-time scaling – actually the ratio between the space and time size – varies according to the size of the solution itself, typically the modulus of u or of Du, according to the kind of regularity under consideration. The reason why the method of intrinsic scaling is necessary can be easily understood by means of simple homogeneity consideration: problems as (1.5) are not scaling invariant in the sense that multiplying a solution by a constant does not yield a solution of another, similar problem. Therefore the resulting lack of homogeneity of local estimates obtainable in the usual parabolic cylinders does not allow for the typical elliptic iterations, and must be therefore re-balanced in some way. The idea is now to pass from inhomogeneous integral inequalities to homogeneous ones involving integrals with solution-dependent supports. To outline how such an **intrinsic approach** works, let us consider a zone, actually a cylinder Q, where, roughly speaking, the size of the gradient is approximately λ – possibly in some integral averaged sense – i.e.

$$|Du| \approx \lambda > 0\,. \tag{1.8}$$

In this case we shall consider cylinders of the type

$$Q = Q_\varrho^{(\lambda)}(z_0) \equiv B_\varrho(x_0) \times (t_0 - \lambda^{2-p}\varrho^2, t_0 + \lambda^{2-p}\varrho^2)\,, \tag{1.9}$$

where $B_\varrho(x_0) \subseteq \mathbb{R}^n$ is the usual Euclidean ball centered at x_0 and with radius $\varrho > 0$ and $z_0 = (x_0, t_0)$. Then cylinders of this type are just the balls with respect to the metric

$$d_\lambda\big((x,t),(y,s)\big) := \max\left\{|x-y|, \sqrt{\lambda^{p-2}|t-s|}\right\}\,. \tag{1.10}$$

Note, when $\lambda \equiv 1$, the cylinder and the metric in (1.9) and (1.10) reduce to the usual parabolic cylinder given by

$$Q_\varrho(z_0) \equiv Q_\varrho^{(1)}(z_0) \equiv B_\varrho(x_0) \times (t_0 - \varrho^2, t_0 + \varrho^2)\,, \tag{1.11}$$

and the corresponding standard parabolic metric defined by

$$d_{\mathcal{P}}\big((x,t),(y,s)\big) := \max\left\{|x-y|, \sqrt{|t-s|}\right\}\,.$$

Indeed, the case $p = 2$ is the only one admitting a non-intrinsic scaling and for which local estimates have a natural homogeneous character. In this case the systems in question are non-degenerate; while in this paper we shall not be interested in the case $p = 2$, treated at length in other parts of the literature, see for instance [**27, 6, 9**] and related references. For the sake of exposition we shall several times restrict to the situation where $p \neq 2$, although several of the arguments proposed here can be easily extended to the case $p = 2$. The **heuristics of the intrinsic scaling method** can now be easily described as follows: assuming that in a cylinder Q as in (1.9) the size of the gradient is approximately λ, we have that the system in (1.5) looks like

$$\partial_t u = \operatorname{div}(\lambda^{p-2} Du)$$

which after a scaling, that is considering

$$B_1(0) \times (-1,1) \ni (x,t) \mapsto v(x,t) := u\big(x_0 + \varrho x, t_0 + \lambda^{p-2}\varrho^2 t\big)\,,$$

behaves exactly as the heat system

$$\partial_t v = \triangle v \qquad \text{in} \quad B_1(0) \times (-1,1)\,,$$

which admits in fact perfect a priori estimates for solutions. The success of this strategy is therefore linked to a rigorous construction of such cylinders in the context of intrinsic definitions. Indeed, the way to express a condition as (1.8) is typically in an averaged sense like for instance

$$\left(\frac{1}{|Q_\varrho^{(\lambda)}|}\int_{Q_\varrho^{(\lambda)}} |Du|^p\, dz\right)^{\frac{1}{p}} \equiv \left(\fint_{Q_\varrho^{(\lambda)}} |Du|^p\, dz\right)^{\frac{1}{p}} \approx \lambda\,, \qquad z \equiv (x,t) \tag{1.12}$$

or

$$\left| \fint_{Q_\varrho^{(\lambda)}} Du\, dz \right| \approx \lambda \,.$$

The problematic aspect in (1.12) clearly relies in the fact that the value of the integral average must be comparable to a constant which is in turn involved in the construction of its support $Q_\varrho^{(\lambda)} \equiv Q_\varrho^{(\lambda)}(z_0)$, exactly according to (1.9).

Let us now briefly turn to the case of standard non-degenerate systems with linear growth i.e. $p = 2$, to recall the known approaches of **partial regularity and local linearization.** On the side of classical partial regularity proofs, the main technique usually employed is basically a linearization one. The basic idea can be now summarized as follows: A point $z_0 \in \Omega_T$ is by definition regular iff the oscillations of the gradient of the solution are small in a quantifiable way in a neighborhood of it. Vice versa, the viewpoint of partial regularity is that this situation is achieved provided the oscillations of the gradient are a priori small in a neighborhood of the point in question, this smallness being measured, as usual, in an averaged integral way. Indeed, functionals as the mean square deviation of the gradient with respect to its average are useful at this stage to express a small oscillation property. The basic assertion of partial regularity is now that a point z_0 is regular iff a smallness condition of the type

$$\fint_Q |Du - (Du)_Q|^2\, dz \le \varepsilon\,, \quad \text{with } (Du)_Q \equiv \fint_Q Du\, dz\,, \tag{1.13}$$

is satisfied for a standard cylinder $Q = Q_\varrho(z_0)$ centered at z_0. Here the number ε implying the regularity of the gradient in general depends on the structure conditions imposed on the system, and in most of the cases also on the point z_0 where Q is centered. Condition (1.13) is in turn used to implement a comparison argument aimed at comparing the original solution u to the solution v of a linear parabolic system with constant coefficients of the type

$$\partial_t v = \operatorname{div}\big(Da\big((Du)_Q\big)Dv\big)\,, \tag{1.14}$$

with v agreeing with u on the parabolic boundary of Q, or at least with v close in an integral sense to u. The role of a smallness assumption as (1.12) is then to quantify the closeness of the original solution to an affine map whose coefficient is given by $(Du)_Q$, so that the system (1.14) can be considered as a Taylor approximation of the original system. On the other hand, since (1.14) is a linear system, good regularity estimates are available for the solution v and in series such estimates can be conveyed to u, thereby proving that Du is Hölder continuous in a neighborhood of the center z_0 of Q. It is of course at the core of partial regularity to prove that a smallness condition formulated in an integral way as in (1.13) is sufficient to make the whole machinery work and to prove the Hölder continuity of Du. This is the standard approach in elliptic and parabolic regularity theory: to commute integral bounds on the oscillation of Du in L^2 – or something near it – in pointwise bounds, that is in L^∞. Needless to say, since a condition of the type (1.13) is only satisfied almost everywhere, the above techniques ultimately lead to almost everywhere regularity in the sense of (1.1).

The first of the methods outlined in the preceding lines, i.e. the intrinsic scaling method, is, as already mentioned above, at the core of DiBenedetto's viewpoint on parabolic regularity and allows for the proof of interior regularity of solutions to systems as (1.5), while the linearization method used to prove partial regularity goes back to the classical papers of DeGiorgi [**14**], and Morrey [**49**] as far as the non-parametric case is considered.

The aim of this paper is now to go beyond; in fact, building on both the approaches, and especially on DiBenedetto's one, we treat the case of regularity for general parabolic systems of the type (1.2) which turn out to feature a degenerate diffusion. This means that an assumption of the type (1.3), usually employed in the literature, is here no longer considered. Needless to say, at the same time, we shall not assume a quasidiagonal structure as (1.6). The only assumption we shall make on the kind of degeneration, which is necessary to quantify the rate of parabolicity of the systems in question, is that the operator $a(\cdot)$ degenerates as the p-Laplacian at the origin, see (1.17) below, therefore a priori prescribing the kind of degeneration at the origin. In other words, the ultimate goal of this paper is to produce methods to match the partial regularity theory available for the case of non-degenerate elliptic and parabolic problems with the degenerate techniques available for the evolutionary p-Laplacian type systems, finding a way to combine the techniques of intrinsic scaling and partial regularity, both outlined above. To achieve this we introduce a number of new tools and methods devised to combine these two approaches, in particular the classical linearization and comparison arguments typical of partial regularity proofs as those using excess functionals as in (1.13), will be implemented in the context of DiBenedetto's intrinsic geometry, eventually leading to a very delicate and technically challenging interplay.

1.2. The main regularity theorem and technical novelties

The specific assumptions we are considering are now listed as follows. Throughout the paper we consider degenerate parabolic systems of the type

$$\partial_t u = \operatorname{div} a(Du) \qquad \text{in } \Omega_T, \tag{1.15}$$

where $a\colon \mathbb{R}^{Nn} \to \mathbb{R}^{Nn}$ is a continuous vector field such that $a \in C^1(\mathbb{R}^{Nn}, \mathbb{R}^{Nn})$ if $p > 2$ and $a \in C^1(\mathbb{R}^{Nn} \setminus \{0\}, \mathbb{R}^{Nn})$ if $p < 2$, satisfying the p-growth condition

$$|a(q)| \leq L\,(1 + |q|)^{p-1}, \tag{1.16}$$

for any $q \in \mathbb{R}^{Nn}$ and some $L \geq 1$.

REMARK 1.1. We remark that in this paper – unless otherwise explicitly stated – we shall not consider the case $p = 2$ which falls into the realm of non-degenerate problems and has been already treated – actually under more general assumptions – in the paper [**27**], to which we refer for results and techniques.

We assume that the vector-field $a(\cdot)$ admits a p-Laplacian type behavior at the origin in the sense that the limit relation

$$\lim_{s \downarrow 0} \frac{a(sq)}{s^{p-1}} = |q|^{p-2} q \tag{1.17}$$

holds uniformly in $\{q \in \mathbb{R}^{Nn} : |q| = 1\}$. Moreover, $a(\cdot)$ is assumed to be **strictly quasi-monotone**, i.e. there exists a constant $0 < \nu \leq 1$ such that for all $q \in \mathbb{R}^{Nn}$ and $\varphi \in C_0^\infty(B_1, \mathbb{R}^N)$ there holds

$$\nu \int_{B_1} \left(|q|^2 + |D\varphi|^2\right)^{\frac{p-2}{2}} |D\varphi|^2 \, dx \leq \int_{B_1} a(q + D\varphi) \cdot D\varphi \, dx\,, \tag{1.18}$$

a condition that can be easily seen to be equivalent to

$$\nu \int_{Q_1} \left(|q|^2 + |D\varphi|^2\right)^{\frac{p-2}{2}} |D\varphi|^2 \, dz \le \int_{Q_1} a(q + D\varphi) \cdot D\varphi \, dz \,,$$

whenever $\varphi \in C_0^\infty(Q_1, \mathbb{R}^N)$. Conditions (1.17) and (1.18) roughly speaking describe the ellipticity properties of the vector field $a(\cdot)$. The first one serves to specify the rate of degeneration at zero – that is the only point where the operator turns out to be degenerate – and it says that the operator degeneration is of the type of the p-Laplacian. Assumption (1.18) is a way to prescribe the ellipticity at those matrices which are different from zero. For instance, whenever a is assumed to be monotone – an additional stronger assumption that we do not actually need in this paper – the following ellipticity condition is implied:

$$c(\nu)|q|^{p-2}|\xi|^2 \le \langle Da(q)\xi, \xi\rangle \qquad \text{for every } q, \xi \in \mathbb{R}^{Nn} \,,$$

which is a degenerate form of (1.3), and tells that the system in question is non-degenerate only when $q \neq 0$.

On the gradient of the vector-field $a(\cdot)$ we do not impose any uniform growth or continuity-condition. We shall merely assume that for given $M > 0$ there exists $\kappa_M \ge 0$ such that

$$|Da(q)| \le L\, \kappa_M |q|^{p-2}, \tag{1.19}$$

for any $q \in \mathbb{R}^{Nn}$ such that $|q| \le M$ and $|q| \neq 0$ if $p < 2$, and a non-decreasing modulus of continuity $\omega_M \colon [0,\infty) \to [0,\infty)$ with

$$\lim_{s \downarrow 0} \omega_M(s) = 0$$

such that $\omega_M^2(\cdot)$ is concave and

$$|Da(q) - Da(\tilde q)| \le \begin{cases} L\,\omega_M\left(\frac{|q-\tilde q|^2}{|q|^2+|\tilde q|^2}\right)\left(|q|^2+|\tilde q|^2\right)^{\frac{p-2}{2}} & \text{for } p > 2 \\ L\,\omega_M\left(\frac{|q-\tilde q|^2}{|q|^2+|\tilde q|^2}\right)\left(\frac{|q|^2+|\tilde q|^2}{|q|^2|\tilde q|^2}\right)^{\frac{2-p}{2}} & \text{for } p < 2 \end{cases} \tag{1.20}$$

whenever $q, \tilde q \in \mathbb{R}^{Nn}$ such that $0 < |q|, |\tilde q| \le M$. The last two assumptions are rather standard in the regularity theory of vectorial elliptic and parabolic problems and important in order to perform the basic linearization arguments when starting from the non-degenerate cases; see for instance [**26, 31, 56**] and related references.

The notion of (weak) solution adopted here, and in the rest of the paper, is of course the usual distributional one and prescribes of course that a map

$$u \in C^0\big(0,T; L^2(\Omega, \mathbb{R}^N)\big) \cap L^p\big(0,T; W^{1,p}(\Omega, \mathbb{R}^N)\big) \tag{1.21}$$

is a (weak) solution to (1.15), under the assumption (1.16), iff

$$\int_{\Omega_T} \big(u \cdot \partial_t \varphi - a(Du) \cdot D\varphi\big)\, dz = 0$$

holds for every $\varphi \in C_0^\infty(\Omega_T, \mathbb{R}^N)$. We now can state our partial regularity result.

THEOREM 1.2 (Main regularity result). *Let*

$$\tfrac{2n}{n+2} < p \neq 2 \tag{1.22}$$

and

$$u \in C^0\big(0,T; L^2(\Omega, \mathbb{R}^N)\big) \cap L^p\big(0,T; W^{1,p}(\Omega, \mathbb{R}^N)\big)$$

be a weak solution of the parabolic system (1.15) *under the assumptions* (1.16) – (1.20). *Then, there exist* $\alpha \equiv \alpha(n, N, p, \nu, L, \kappa_3) \in (0,1)$ *and an open subset* $Q_0 \subseteq \Omega_T$ *such that*

$$Du \in C^{\alpha,\alpha/2}_{\mathrm{loc}}(Q_0, \mathbb{R}^{Nn}) \qquad \text{and} \qquad |\Omega_T \setminus Q_0| = 0\,.$$

In other words the previous result shows that in a neighborhood of a point $z_0 \in Q_0$ the spatial derivative Du is Hölder continuous with Hölder-exponent α with respect to the standard parabolic metric defined in (1.11). This result extends the classical partial regularity available for non-degenerate systems to those admitting a degeneration of p-Laplacian type at the origin.

REMARK 1.3 (On condition (1.22)). We remark that the lower bound on p appearing in (1.22) is absolutely natural to prove regularity in such a context; see for instance [**17**] for a larger discussion and conterexamples. The assumption $p \neq 2$ is actually non-restrictive. In Remark 1.1 we already mentioned that the case $p = 2$ turns automatically out to be non-degenerate and a partial regularity result can be obtained under weaker assumptions; for results and techniques in this case we refer to the paper [**27**].

As mentioned in the previous section, the proof of Theorem 1.2 is based on a very delicate interaction between the two approaches, i.e. the intrinsic scaling method and partial regularity and local linearization. The Hölder continuity of the spatial gradient Du around a (regular) point z_0 is achieved via a suitable decay of an **excess-functional** $\Phi(\cdot)$ of the type appearing in (1.13) and which measures in an integral way the oscillations of the gradient Du, being simultaneously linked to the structure of the problem under consideration. In fact, in this case $\Phi(\cdot)$ takes a more peculiar form. For $p > 2$ – a case which in this introduction we restrict to for clearness of exposition – $v \in L^p(0,T; W^{1,p}(\Omega, \mathbb{R}^N))$ and $Q_\varrho(z_0) \subseteq \Omega_T$ we set

$$\Phi(v; z_0, \varrho) := -\!\!\!\!\!\!\int_{Q_\varrho(z_0)} \left(|(Dv)_{Q_\varrho(z_0)}|^2 + |Dv - (Dv)_{Q_\varrho(z_0)}|^2\right)^{\frac{p-2}{2}} \cdot |Dv - (Dv)_{Q_\varrho(z_0)}|^2 \, dz, \tag{1.23}$$

where $(Dv)_{Q_\varrho(z_0)}$ denotes the mean value of Dv on $Q_\varrho(z_0)$. Here $Q_\varrho(z_0)$ denotes the standard parabolic cylinders defined in (1.11).

The ultimate goal is to prove a decay estimate of the type

$$\Phi(u; z_0, \varrho) \le c\, \varrho^\alpha \tag{1.24}$$

at the regular point z_0 and eventually in a neighborhood of z_0, an inequality which by mean of a standard integral characterization of Hölder continuity due to Campanato and Da Prato implies the Hölder continuity of the Du. In turn, in order to apply the intrinsic geometry approach we shall define an intermediate excess functional incorporating an auxiliary scaling parameter $\lambda > 0$; this is defined by

$$\Phi_\lambda(v; z_0, \varrho) := -\!\!\!\!\!\!\int_{Q^{(\lambda)}_\varrho(z_0)} \left(|(Dv)_{Q^{(\lambda)}_\varrho(z_0)}|^2 + |Dv - (Dv)_{Q^{(\lambda)}_\varrho(z_0)}|^2\right)^{\frac{p-2}{2}} \cdot |Dv - (Dv)_{Q^{(\lambda)}_\varrho(z_0)}|^2 \, dz, \tag{1.25}$$

where $Q^{(\lambda)}_\varrho(z_0)$ is defined in (1.9) and $(Dv)_{Q^{(\lambda)}_\varrho(z_0)}$ denotes the mean value of Dv on $Q^{(\lambda)}_\varrho(z_0)$; that is, we are using an excess functional defined on **intrinsic cylinders** relating the scaling λ of the domain of integration to the gradient Du in a way described in (1.12). Note that $\Phi_\lambda(v; z_0, \varrho)$ with $\lambda \equiv 1$ reduces to the excess functional $\Phi(v; z_0, \varrho)$ from (1.23), and as it will be clear from the proof, this excess functional will be eventually linked to the one in (1.23) during the proof, when controlling the size of the gradient during certain iteration schemes. Specifically, at every scale the value of the number λ will be related to the solution via an intrinsic relation of the type (1.12). In a second step we show that the values of the numbers λ stay bounded as $\varrho \to 0$ provided we are in a neighborhood of a regular point. Furthermore, if $Du(z_0) = 0$ then λ goes to zero, when $\varrho \to 0$. Therefore the intrinsic – i.e. stretched in time – cylinders $Q_\varrho^{(\lambda)}(z_0)$ will be ultimately comparable to the standard parabolic ones in a neighborhood of a regular point z_0 with $Du(z_0) \neq 0$. This features a new technical approach and means that in such zones the two excess functionals in (1.23) and (1.25) will be finally comparable.

At this point it is rather clear that the interaction between the role of the intrinsic geometry – which appears since the problem is degenerate – and the standard linearization methods is the crucial point of the proof of Theorem 1.2, since the shape of the cylinders, and therefore of the estimates involved, change according to the local degeneration rate of the system. A closer look to the strategy of the proof, which, using different comparison techniques according to whether we are in a degenerate point or in a non-degenerate one – a common alternative occurring in degenerate problems, see for instance **[26, 56]** – presents outstanding technical difficulties. The strategy is roughly to distinguish between two regimes.

The non-degenerate regime (NDR). We consider an intrinsic cylinder $Q_\varrho^{(\lambda)}(z_0)$ and we say we are in the non-degenerate regime when the average of the gradient is larger than the excess functional (whose value has to be understood in this context as a re-normalization factor in the local linearization procedure):

$$\Phi_\lambda(u; z_0, \varrho) \ll |(Du)_{Q_\varrho^{(\lambda)}(z_0)}|^p .$$

Here, λ is again coupled to the gradient in an intrinsic way according to a relation as (1.12). More precisely, we assume that

$$|(Du)_{Q_\varrho^{(\lambda)}(z_0)}| \approx \lambda .$$

This last relation actually makes $Q_\varrho^{(\lambda)}(z_0)$ an intrinsic cylinder. In this situation we adopt the local linearization procedure from the partial regularity theory and locally compare the solution u with a solution v of a linear parabolic system with constant coefficients, i.e. a problem which is non-degenerate; this is achieved via the method of $\mathcal{A}$-caloric approximation from **[29, 54]**. Then good a priori estimates for v are indeed inherited by u in the regularization process. For this we refer to Section 8.1.

The degenerate regime (DR). This case refers to the situation when the gradient is in average smaller than the excess functional in the sense that

$$|(Du)_{Q_\varrho^{(\lambda)}(z_0)}|^p \ll \Phi_\lambda(u; z_0, \varrho) \quad \text{and} \quad |(Du)_{Q_\varrho^{(\lambda)}(z_0)}| \le c\lambda ,$$

or

$$|(Du)_{Q_\varrho^{(\lambda)}(z_0)}| \ll \lambda .$$

This situation is considered in Section 8.2 and involves the use of one of the main tools developed in this paper, the *p-caloric approximation lemma*. In this situation we compare the original solution u with the solution w of a degenerate parabolic system of p-Laplacian type

$$\partial_t w = \operatorname{div}(c|Dw|^{p-2}Dw) \tag{1.26}$$

and then proceed taking advantage of the fact that also in this case the solution w is regular and enjoys certain a priori estimates, eventually inherited by u. It is precisely this point where we exploit the fact that we are using intrinsic cylinders: indeed only the use of such a geometry makes a comparison between u and w possible since alone on such cylinders regularity properties of w can be expressed in a way that allows them to be transferred to u. At this stage essentially the method of intrinsic geometry described in the previous section comes into the play. In this context it is an important part of the proof to put the original estimates of DiBenedetto & Friedman in ready-to-use form; see Chapter 7. We note that the closeness of the original solution u to w is a consequence of the assumption (1.17), which prescribes the type of degeneration of the original parabolic system at the origin; once used in a suitable linearization scheme assumption (1.17) reads as a closeness condition between u and the solution to an asymptotic system as (1.26).

The linearization methods corresponding to the non-degenerate regime (NDR) and the degenerate regime (DR) are contained in any case in Chapter 6, where the perturbation lemmas necessary to the use of the $\mathcal{A}$-caloric and p-caloric approximation lemmas are reported.

Finally, the degenerate and the non-degenerate regime can be matched via an extremely delicate iteration procedure where keeping the control of the constants is by no means a trivial fact; see Section 8.3 and the tables with the constant dependencies can be found in Remark 8.4.

The scheme of proof outlined above also yields a precise characterization of the singular set which is displayed in the next theorem.

THEOREM 1.4 (Description of the singular set). *Under the assumptions of Theorem 1.2, we know that the singular set* $\Sigma = \Omega_T \setminus Q_0$ *is contained in* $\Sigma_1 \cup \Sigma_2$, *where*

$$\Sigma_1 \equiv \left\{ z_0 \in \Omega_T : \liminf_{\varrho \downarrow 0} \Phi(u; z_0, \varrho) > 0 \right\},$$

$$\Sigma_2 \equiv \left\{ z_0 \in \Omega_T : \limsup_{\varrho \downarrow 0} |(Du)_{Q_\varrho(z_0)}| = \infty \right\}.$$

Moreover, if for some regular point $z_0 \in \Omega_T \setminus (\Sigma_1 \cup \Sigma_2)$ *there holds* $Du(z_0) \neq 0$, *then there exists* $\sigma > 0$ *such that* $Du \in C^{\alpha,\alpha/2}(Q_\sigma(z_0), \mathbb{R}^{Nn})$ *for any* $\alpha \in (0,1)$.

1.3. The p-caloric approximation technique

As mentioned above, one of the aims of this paper is to introduce a compactness technique for treating the regularity and the linearization of degenerate parabolic systems as in (1.15). Such compactness methods are a powerful tool in the modern theory of partial differential equations in that their use allows to simplify approaches and proofs, and to often achieve optimal regularity results unreachable otherwise. Here, by compactness methods we mean the use of *convergence methods in order to prove certain inequalities*, which, in principle, could also be proved by direct, analytical arguments. Such analytical methods are very often delicate, and do not always lead to the optimal result one would expect; this is one of the reasons for using instead indirect methods. A basic, elliptic example of such methods is given by the following:

THEOREM 1.5 (p-harmonic approximation lemma [**25**]). *Let $n, N \in \mathbb{N}$ with $n \geq 2$ and B be the unit ball in $\mathbb{R}^n$ and $p > 1$. For every $\varepsilon > 0$ there exists a positive constant $\delta_0 \in (0,1]$ depending only on n, N, p and ε such that the following is true: Whenever $u \in W^{1,p}(B, \mathbb{R}^N)$ satisfying $\|Du\|_{L^p(B)} \leq 1$ is approximately p-harmonic in the sense that*

$$\left| \int_B |Du|^{p-2} Du \cdot D\varphi \, dx \right| \leq \delta_0 \sup_B |D\varphi| \tag{1.27}$$

holds for all $\varphi \in C_0^\infty(B, \mathbb{R}^N)$, then there exists a map $h \in W^{1,p}(B, \mathbb{R}^N)$, such that $\operatorname{div}\,(|Dh|^{p-2} Dh) = 0$ *in B, and such that*

$$\int_B |Dh|^p \, dx \leq 1 \qquad \text{and} \qquad \int_B |h - u|^p \, dx \leq \varepsilon^p \,. \tag{1.28}$$

The case $p = 2$ of Theorem 1.5 has been used by DeGiorgi in his fundamental paper on the regularity of minimal surfaces [**15**] and eventually by Simon [**58**] to study the regularity of harmonic maps, while a more general version – the $\mathcal{A}$-harmonic approximation method – has been introduced by Duzaar & Steffen [**30**] in the setting of Geometric Measure Theory; this was later on applied in the setting of elliptic systems in [**23**] to yield an optimal regularity result. The case $p \neq 2$ has been proved in [**25**] and was used to establish regularity results for various degenerate problems [**25, 26**]. The difficulty in the case $p \neq 2$ obviously relies in passing to the limit in the vector field, since the vector field involved is non-linear, and plain weak convergence arguments do not suffice.

A first parabolic extension – suitable for non degenerate problems and involving linear operators – has been given in the papers [**27, 29, 54, 6**], while a general non-linear parabolic version of Theorem 1.5 – together with its application to the regularity of degenerate parabolic problems – was still lacking; we refer to the survey paper [**28**] for an updated overview on compactness methods.

The second main result of the paper is in fact concerned with this last gap, thereby providing a suitable parabolic analog of Theorem 1.5, that we indeed call the p-caloric approximation lemma, see Theorem 1.6 below. This result indeed provides an analog which allows to approximate solutions of the parabolic p-Laplacian system – or even more general degenerate parabolic systems – with exact solutions of the system, and therefore to perform the local linearization methods necessary to prove partial regularity as explained in the previous section.

For the sake of clearness, in this introductory part we report a version of the p-caloric approximation lemma for the model case given by parabolic p-Laplacian. However, our methods also allow to treat more general vector-fields $\mathcal{A}(z, w) \approx |w|^{p-2}w$, without any further effort and therefore we prove the result for general vector fields with a p-Laplacian structure at the origin, also allowing a dependency on the variable $z = (x, t)$. For the precise structure conditions on $\mathcal{A}$ we refer the reader to Chapter 4. At this stage we state the p-caloric approximation for the degenerate case $p \geq 2$ only, again for the sake of clearness. Later on we also provide a version for the singular case $2n/(n+2) < p < 2$, whose proof needs certain adjustments compared to the case $p \geq 2$. This can be found in Section 4.2, Theorem 4.5. For the notation used in the following we refer to Section 2.1 below.