FIVE GAPS IN MATHEMATICS

A.Bahri

0.Introduction.

In the following, we describe five proofs or methods devised for landmark results in Mathematics, since 1984, which we find would require further precisions, the filling up of details or corrections. The gaps range from the clear mistake to the existence of an important point that is not obvious to us.

We leave it to the reader to decide the nature of the gaps that we describe.

This work is the result of a long period of reading completed by the author, with the help and support of several discussions, mainly email conversations, with various colleagues and friends. This includes also several authors of the papers that we will discuss.

We would like to thank here all these friends and colleagues who took the time to respond to our queries. This has been a genuine process of learning for the author.

The field of Mathematics has undergone a considerable acceleration since 1984, with the increased number of mathematicians and the transformation of the techniques and methods of proof. A certain amount of speculation, or unfinished proofs has been tolerated over the last thirty years, not uniformly, but at least partially. Some gaps have appeared as a consequence of this change in the standards of proof; they sometimes require adjustments that go much beyond partial modifications: fundamental issues have for example been ignored at least partially in the area of pseudo-holomorphic curves for contact structures, they have been ignored, but they corresponded to very real difficulties. Hopefully, this will be adjusted and the right framework and results will be found.

We see the present paper as a contribution to this complicated process of correction. We hope that we have no mistake ourselves in our assessment[1]and we hope that we are not offending any colleague: we know that, behind every mathematical paper, there is a considerable amount of work.

1.S.K.Donaldson's proof of his celebrated theorem.

S.K.Donaldson's original proof of his celebrated theorem [27] is not complete according to us: in order to solve the situation near a reducible Yang-Mills connection and in order to find that the moduli space of self-dual connections defines a cone over CP^2, a perturbation is introduced in [27] in the chart where the self-dual Yang-Mills equations, which are nonlinear equations, are reduced to the addition of a Fredholm linear operator $T = T(x)$ and a finite dimensional map $\phi = \phi(z)$; z lives in a finite dimensional space. ϕ is perturbed into $\tilde{\phi}$, also valued in the same finite dimensional space than ϕ; this in order to create the conical CP^2 at the reducible. The precise formula for the perturbation $\tilde{\phi}$ reads, see [27]: $T(x) + \phi(z) + \eta\beta(\frac{|z|}{\eta})Lz$, z is in $kerT$, whereas x is in appropriate infinite dimensional complement space. L is an S^1 equivariant linear map from C^{p+3} to C^p. This technique is later extended to neighborhoods of irreducible connections.

[1]In an earlier version of the present paper, we had stated that the original proof of S.K.Donaldson [27] celebrated theorem and the proof by D.Freed and K.Uhlenbeck [29] both contained gaps. This was not accurate and our understanding has evolved. Also, after this note was included in our personal website, Professor Ernst Kuwert, from the University of Freiburg, Germany, has informed us that the solution of the graph issue that we raise in the proof of the positive mass conjecture by R.Schoen in [50] is written in his thesis. This thesis is unfortunately not widely referred to in the literature; see section 2 of the present note for a discussion of this graph issue as we undertook it, without the information about the work of Professor Kuwert-as we went from one paper to another one, from an email to another one, trying to find a proof-and also for the idea of the proof of this point, which we outline in a footnote.

A basic tool for this perturbation technique is the definition of "slices": these are, near a Yang-Mills or even near a more general connection A, a set of connections $A+a$ defined by the equation $d_A^*a = 0, |a|_{W^{3,2}} \lneq \epsilon$. A slice defines, at an irreducible connection A, a section to the gauge action, see Theorem 6 of [27]. At a reducible connection, a slice is a section to the quotient of the gauge group by S^1, S^1 fixing the reducible connection.

Because of the Gribov ambiguity [58], there is a need to have several different slices. The local formulations of the self-dual Yang-Mills equations change as we change slices, see [27], [29].

Slices are a priori gauge-dependent in their definition since their construction requires the choice of a base connection along a given gauge-orbit and since it is not obvious that a slice remains a slice after gauge transformation. This gauge-dependency complicates the discussion as one cannot speak in an intrinsic way about these slices, their distances etc.

Similarly, two slices can be completely disconnected, whereas their images through the projection map p of [27] intersect. This means that these two slices intersect after modifying one of them with the action of a gauge-group element or after redefining this slice at another base point of the same gauge orbit: the gauge group G interferes with the argument.

Would the gauge group G be a compact Lie group, then these difficulties could be dealt with. We could average over the action of G and define slices in a G-invariant way, distances of slices etc. But G is non-compact and infinite dimensional.

In order to simplify the presentation of the issues that we think we have discovered in [27], we will sidestep this problem here.

It is proven in Theorem 6 of [27] that slices (we are discussing mainly irreducible slices in the sequel) as defined above define local sections to the gauge group action, this follows from the use of the implicit function theorem.

These sections are, in principle, only local; that is, given a in the slice of $A(d_A^*a = 0, |a|_{W^{3,2}} \leq \epsilon)$, given g an element of the gauge group and denoting $g.(A+a)$ the action of the gauge element g on the connection $(A+a)$, assume that $g.(A+a) - A$ is in the same slice than a and assume that $|g - Id|_{W4,2}$ is appropriately small; then we can conclude that $g = Id$. This is a corollary of Theorem 6 of [27] and this holds true.

On the other hand, if the condition $|g - Id|_{W^{4,2}}$ small is not verified, an additional argument is needed -this happens all the time in dynamical systems, where local sections are not to be confounded with global ones-in order to prove if $g.(A+a)$ is equal to $A+b$, b in the same slice than a, then $g = Id$.

The values of the self-dual Yang-Mills equations at $A+a$ and at $g.(A+a) = A+b$ are related by the adjoint action [27], [29] of g on the curvature of the connections. This law should be preserved by the perturbation; this condition on the perturbation is neither discussed nor introduced by S.Donaldson in [27], because it is clear to the author of [27] that Theorem 6 of his paper excludes this occurrence. Let us observe at this point that, if this occurrence were to happen, then the related condition on the perturbation would be hard to satisfy. It is therefore important to check carefully that the arguments of [27] imply indeed that the equation $g.(A+a) = A+b$, with g not equal to Id and a and b in the same slice, is never verified.

.**Size of the (irreducible) slices**

We find in what follows that the proof of Theorem 6 of [27] uses a further specification or restriction on the size of the slice relative to the behavior of the operator d_A. This restriction is implicit in [27]. Let us make it explicit: given a slice, how large can $|g - Id|_{W^{4,2}}$ be if the equation $g.(A+a) - A = b$ with $g \neq Id$ is verified? We can understand this by re-reading the proof of Theorem 6 of [27], where we find the equation in the vicinity of an irreducible:

$$d_A g = bg - ga$$

In order to understand where g is valued, we denote, as in [27], P the $SU(2)$-principal bundle over the manifold M of dimension 4. The elements of gauge group G of P may then be viewed as sections of the associated bundle $S = P \underset{\text{adj}}{\times} SU(2)$, where $SU(2)$ acts on itself by conjugation. In this way, g is viewed, in a local trivialization chart for P, as a map from the open set of this chart in M into $SU(2)$, up to conjugation by the adjoint action of the transition functions when this trivialization chart is changed. We denote Z the Lie Algebra of $SU(2)$. Let T denote then the bundle $P \underset{\text{adj}}{\times} Z$ associated to P and defined with the use of Z. $SU(2)$ acts on Z through the adjoint action.

Coming back to the above equation and viewing g as a section of S, we want to derive an estimate on $|(g-Id)|_{W^{4,2}}$ if g is a solution of the above equation. We use then the argument of S.K.Donaldson in the proof of Theorem 6 of [27]; but we first modify this argument and we change it into a slightly more explicit technical argument: We write g as $g = fId + k$; $2f$ is the trace (it is real) of g and k is a section of T. This means that k is also defined in local trivializations of P by maps from the open sets in M of these trivializations into Z; these maps are "glued" through the adjoint action of the transition functions (valued into $SU(2)$) on Z.

Since, in a local trivialization of T, k reads as an element of Z, k verifies, with $k^+ = \bar{k}^t$, $k + k^+ = 0, \mathrm{trace} k = 0$. We denote in the sequel $\Gamma(T)$ the space of sections of T.

The condition that $g \in SU(2)$ reads:

$$f^2 + \frac{|k|^2}{2} = 1$$

The above equation on g rewrites in an intrinsic way (although the writing uses a trivialization, the equations are intrinsic; observe that $a + a^+ = 0, b + b^+ = 0, \mathrm{trace} a = 0, \mathrm{trace} b = 0$, a and b are in $\Omega^1(Z)$, the space of 1-forms of M with values in T):

$$4df = \mathrm{trace}(kb + bk - (ak + ka))$$

$$2d_A k - (bk - kb) + (ka - ak) = 2f(b - a)$$

Defining then δ to be a lower-bound for the ratio $\frac{|d_A w|_{W^{3,2}}}{|w|_{W^{4,2}}}$ for $w \in W^{4,2}(\Gamma(T))$-these are sections of T, with regularity $W^{4,2}$-, we use the L^∞-bounds on f and k and we derive after bootstrapping the estimate:

$$|g - fId|_{W^{4,2}} + |f - [f]|_{W^{4,2}} \leq C\epsilon/\delta$$

where C is an appropriate constant and $[f]$ is the mean value of f. Without loss of generality, we can assume that $[f] \geq 0$. It is clear that $|g - fId|_{W^{4,2}} \leq |g - Id|_{W^{4,2}} \leq |g - [f]Id|_{W^{4,2}} + |1 - [f]|_{L^\infty}$.

The above estimate is the best estimate that we have been able to derive. We have tried to find a better estimate and we have completed some additional work which uses the linearity of the equation on k, assuming also that f is very close to 1 in an L^∞-sense. No significant progress has come out of this attempt, although a much better understanding of the solutions of this equation can be derived when g is close to Id, in an L^∞-sense to the least.

Let us observe here that there is another nonlinear formulation of the above equation. This formulation exploits the fact that $d_A * a = 0$ and reads:

$$d_A * (g^{-1} d_A g) = -d_A * (g^{-1} b g)(*)$$

This is actually also the form that is used below in order to understand the conditions under which the implicit function theorem holds. Under this form, it is quite clear that if $|g - Id|_{L^\infty}$ is small, then:

$$\int |d_A g|^2 \leq C \int |b|^2 (**)$$

This estimate is straightforward. Bootstrapping $(**)$, we derive again the same estimates as above. Some work removes the smallness assumption on $|g - Id|_{L^\infty}$.

As above, we assume, without loss of generality, that, if g reads $(fId + k), k + k^+ = 0, \mathrm{trace} k = 0$, then $[f] \geq 0$.

We study now the conditions under which we can use the implicit function theorem. This theorem tries to reduce an equation to its linearized form and assumes that this linearized form is invertible. The equation is the one written above, with b small in $W^{3,2}$, $d_A * b$ is not necessarily zero anymore. If $(g - Id)$ is small in an L^∞-sense, then this equation becomes close to a Laplace equation with the operator $d_A * d_A$ acting as a Laplacian and a right hand side equal to $-d_A * b$. Such an approximate equation has a unique solution. We can even keep the operator to be nonlinear, of the form $d_A * g^{-1}d_A g$(this nonlinear operator has some nice properties) and we can try to solve this equation, but we need to replace g with Id in the right hand side in order to develop an understanding of the existence and uniqueness of solutions. This leads to a natural minimal assumption in order to find a unique solution for $(*)$: that $|g-Id|_{L^\infty}$ should be small. We claim that this assumption implies then a stronger smallness assumption on $|g-Id|_{W^{4,2}}$.

Indeed, let us assume-this is not needed, but the argument is then easier-after subdividing a slice into a finite, fixed number of smaller slices, that ϵ is small. We consider an irreducible slice which intersects a fixed, small neighborhood of a given self-dual reducible Yang-Mills connection. We can then bootstrap this L^∞ estimate using a decomposition for d_A into a finite dimensional space where the norm of d_A is small-A is close to a reducible Yang-Mills connection A_{red}; all norms are equivalent along this finite dimensional space which is the kernel of $d_{A_{red}}$ and the constants depend only on how close A is to A_{red}- and a supplement; we find that this estimate, up to some conversion constants, is in fact as an estimate on $|g-Id|_{W^{4,2}}$. This is the only general estimate on $|g-Id|_{W^{4,2}}$ from which an application of the implicit function theorem follows. We therefore assume that this estimate holds in the sequel.

We could have skipped the transformation of the L^∞-estimate into a $W^{4,2}$ estimate and we could have assumed this $W^{4,2}$ estimate from the beginning; but we are trying to see how far this can be pushed and a significant result be still derived.

Returning now to the nonlinear equation and considering a general solution g, we know that $|(g-Id)|_{W^{4,2}}$ is larger than $|g-fId|_{W^{4,2}}$. The only estimate that we have on this latter term is that it is of the order of ϵ/δ. It therefore comes naturally to mind to require that the size ϵ of the (irreducible) slice should be $o(\delta)$. Under such a requirement, $|k|_{L^\infty}=o(1)$ and therefore, $1-[f]=o(1)$, so that $|(g-Id)|_{W^{4,2}}=o(1)$. There is more work, but we can then start thinking about applying the implicit function theorem.

This conclusion, viewed as the derivation of an upper-bound on the size of the slices, is quite natural. This estimate implies in turn, after some computations, that, for every B in the same slice, the operator d_B changes by an amount of the order of $o(\delta)$ at most, so that a lower-bound for d_B in the slice is $\delta/2$. This conclusion has been derived under the assumption that $\epsilon=o(\delta)$; but, we know that each time we use the implicit function theorem, then the associated linearized operators (such as d_B) change very little over the domain of application of this theorem.

There are of course more sophisticated versions and uses of this theorem which require further verifications, a detailed study of the reduction to the linearized equation and a careful checking of the invertibility of the linearized operator over the whole domain (the slices will then have to be defined by other equations than $d_A^* a=0, |a|_{W^{3,2}}\lesssim\epsilon$ and they should not contain any reducible connection, since the linearized operator has a kernel at these connections). However, these more sophisticated versions use the fact that the map is then a local embedding and a global immersion. This is of little use on domains (the slices) which have a boundary as some pre-images, as the image runs in the target space, could "escape" through the boundary and their number cannot therefore be deemed to be constant, equal to 1.

It follows then, under such the natural assumption $\epsilon=o(\delta)$, that the order of the size of a slice (up to a fixed constant $C\in[1/2,2]$)can be derived from the computation of a lower-bound for the operator d_B, acting from $W^{4,2}$ to $W^{3,2}$ (these are Sobolev spaces of sections of the associated bundle defined by the Lie Algebra of $SU(2)$, see[27], [29]), and taken at an arbitrary connection B of the slice. This is used below.

.Questions about the construction of the slices

We come back then to the perturbation introduced by S.Donaldson in [27] near a reducible Yang-Mills connection. It reads $Tx+\tilde{\phi}(z)=Tx+\phi(z)+\eta\beta(\frac{|z|}{\eta})Lz$, z in in $kerT$, the kernel of the linearized tangent operator to the Yang-Mills equations at a reducible Yang-Mills solution, x is in the supplement space, the differential of ϕ at zero is zero ($d_0\phi=0$). Clearly, there is a need of irreducible slices at distance of the order η from the reducible Yang-Mills connection A_{red}, because the linearized operator at a solution of the perturbed equations at the distance order of η might not be surjective. Since the operator $d_{A_{red}}$ has a kernel, a lower-bound for the norm of the operator d_A at such a distance is of the order of the distance, which is of the order of η. The size of an irreducible slice should then be $O(\eta)$ to the most, see above and the argument can be repeated to show that an irreducible slice is needed at distance μ, with μ of the order of an integer multiple of η from the reducible, giving a slice of size at most $O(\mu)$

to avoid holonomy. It follows that the perturbation argument cannot use a sequence of values for η tending to zero as we would then need to add constantly new slices, increasing their number to infinity.

Observe also that, since this size depends on the parameter η of the perturbation, the construction of the slices must be completed in function of the perturbation to come (!).

As a matter of fact, it is not clear to us how these slices should be built; an assumption that comes naturally to the mind after reading [27]-see section II.2, p289 of [27]- is that it suffices to have slices (irreducible as above and also one additional slice at each reducible self-dual connection) that define through the projection p of [27] a covering of the set of self-dual Yang-Mills connections. This set, once excluded a neighborhood of its boundary, is compact; so that the construction of this covering is derived nicely; this is very convenient.

However, the perturbed equations, see above, might have singular solutions z_{per} (with non-surjective differential) with $|z_{per}| = O(\eta)$. Indeed, solutions of the perturbed equations solve $\phi(z) + \eta\beta(\frac{|z|}{\eta})Lz = 0$, with $\phi(z) = O(|z|^2), d_z\phi = O(|z|)$. There is an additional condition associated to the non-surjective differential which can be explicitly written, but does not help the estimates much.

Since $d_{z_{per}}\phi = O(|z_{per}|)$ and since $\phi(z_{per}) = -\eta\beta(\frac{|z_{per}|}{\eta})Lz_{per}$ at a solution of the perturbed equation, whereas the set of self-dual connections is locally defined by the equation $\phi(z) = 0$, the only information that we are able to derive from the above equations is that the distance of z_{per} to this set could be $O(\eta)$ or more (this depends on the precise value of ϕ, L, β; concrete examples can be built).

Since the (irreducible) slices are themselves $O(\eta)$, most probably $o(\eta)$, it is now unclear whether it would suffice to take them (with the reducible ones added to them) to cover, through p, the set of self-dual Yang-Mills connections. Indeed, their size is smaller or of the same order than the only available (and accurate as some examples of ϕ etc show) estimate on the distance of some solutions of the perturbed equations near the reducible Yang-Mills connections to the set of self-dual Yang-Mills connections; we might then need to take a larger set of connections and related (new, other) slices since the ones that we are considering so far might not include all the solutions to the perturbed equations.

Furthermore, Fredholm equations are only locally proper. In order to conclude that the moduli spaces remain compact through perturbation, we need to build the slices so that these Fredholm equations-after the initial perturbations as above in the reducible Yang-Mills slices-can be reduced with the use of a single chart within each given slice, at most a finite, fixed number of them, to the form $T + K$ of [27].

This will reduce further the size of the slices and this complicates their construction.

There is also the need to verify that the equations remain Fredholm after the initial perturbations. This is not totally obvious because there is a non constant gauge transformation from slice to slice, which is highly non-linear and forces the new equations to be Fredholm only in a very small neighborhood of the moduli space, if this property is indeed verified (as it seems to be).

In addition, this gauge transformation, from an irreducible slice into a reducible slice, requires the choice of a section to the S^1-action in the reducible slice. The gauge transformations become very local, as well as the Fredholm features and the charts of reduction.

Several important details are missing here and they require attention and a step by step construction.

.**An additional important issue**

We now have recognized that some of the new slices that we have to build are small in size-$o(\eta)$- and are at a distance (after gauge re-normalization possibly) $O(\eta)$ from the reducible Yang-Mills connection. They do not correspond to the slices covering the moduli space of self-dual Yang-Mills connections. As noted above, such slices, after gauge re-normalization (possibly), might be entirely contained inside the slice of the reducible Yang-Mills connection in which the perturbation above has been introduced.

This slice (of the reducible connection) is invariant under the action of an S^1-subgroup of the gauge group. The new equations are Fredholm in the slice; but it is difficult to perturb the equations so that the moduli space becomes a manifold within this slice, because the S^1-invariance of the equations has to hold through the perturbations.

Indeed, the search for regular values has to be completed in sections to the S^1-action. The charts of reduction are local and it is highly unclear, due to holonomy, that the resulting perturbation will be well-defined as an S^1-covariant map: starting from some connection in the slice, after completing some loop in order to solve the singularities of the moduli space, we might come back to the same connection, shifted by the action of an element of the gauge sub-group S^1 that acts inside the slice, with a corresponding shift in the section and with a corresponding shift in the equations. However, the regular value found at the initial connection and the regular value derived at the end of the loop-the "limit" of the process along the loop-might not correspond one to the other through the action of this element of the gauge sub-group S^1.

We might then think of a process where we would mod out by the action of S^1-viewed as a sub-group of gauge transformations-near the zero set of the new equations (after the perturbation in the reducible slices) in the slice of the reducible self-dual Yang-Mills connection. We would also mod out out the target space $W^{2,2}$ of $\Omega^2_-(Z)$-the space of two forms on M valued in T, with regularity $W^{2,2}$- by the adjoint action of this sub-group of gauge transformations. But the two-form ω_0 in $\Omega^2_-(Z)$, which is identically equal to 0 at every point of the manifold M, is the value for the equations and the value for the corresponding map which we want to perturb using the transversality theorem; this is a fixed point for the adjoint action and this action has in fact many more fixed points in $\Omega^2_-(Z)$. The quotient space is very singular and the solutions of the equations correspond to the pre-images through the associated equivariant map (eg Φ, p289 of [27] and its perturbations) of one of these singularities. This seems difficult to achieve and is certainly not classical as all transversality theorems are based on the regularity of the target spaces.

Completing the search for an appropriate perturbation with the use of the irreducible slices that are transverse to this S^1-action apparently solves this problem. However, we have discovered that it is very much possible that many of the new slices that we build in order to solve the singularities of the moduli space within the reducible slice are in fact contained, after gauge transformation, inside this reducible slice. They may be seen then, after appropriate-and non obvious- gauge transformation, as sections to the S^1-action in the reducible slice.

After this observation, the process of finding an appropriate perturbation for the new equations with the use of these slices becomes equivalent, in certain cases to the least, to the same process within the reducible slice.

We know that the latter cannot be performed in full generality, due to the S^1-invariance that needs to hold through the perturbation. The argument of [27] cannot proceed then and the proof of S.K.Donaldson is not complete according to us.

Let us observe here that there is another proof of S.K.Donaldson's celebrated theorem [29] by D.Freed and K.Uhlenbeck. Two proofs for the understanding of the behavior of the moduli space at a reducible self-dual connection are proposed in [29]. Theorem 4.11 of [29]does not allow to conclude the proof of S.K.Donaldson's theorem: new singularities of the moduli space may be created, as some explicit examples show, by the perturbation proposed in this Theorem.

On the other hand, the proof of Theorem 4.19 of [29], based on a metric perturbation argument, is complete according to our understanding. Therefore, the proof of the celebrated theorem of S.K.Donaldson is complete following the arguments of [29].

2. The Positive Mass Conjecture in dimensions 4 to 7; the Yamabe conjecture in dimensions 4 and 5 (Non Conformally Flat).

R.Schoen [47] published in 1984 a paper in Journal of Differential Geometry where he established a direct relation between the proof of the positive mass conjecture and the proof that there is a solution, which is an absolute minimum of the corresponding variational problem, to the Yamabe conjecture.

R.Schoen and S.T.Yau [55] had published in 1979 a paper about the positive mass conjecture in dimension 3, see also the more recent work by T.Colding and W.Minicozzi II [22] in order to understand the behavior of the minimal surface Σ at infinity; E.Witten [60] established this conjecture for all spin manifolds under a "dominant energy condition", see the work by T.Parker and C.H.Taubes [42] to clarify the meaning of this condition, so that the proof of the Yamabe conjecture, using the technique proposed by R.Schoen in [48], is complete for all three dimensional (compact) manifolds and more generally for all spin (compact)manifolds under some additional "dominant energy" condition.

R.Schoen and S.T.Yau announced also in [56] the proof of this conjecture for all conformally flat manifolds. T.Aubin [6] had proven the Yamabe conjecture for all non-conformally flat (compact) manifolds of dimension $n \geq 6$. This gives a large class of manifolds for which we do know, according to all these references, that the solution of the

Yamabe conjecture exists and is an absolute minimum. What is left from this list is the case of non conformally flat (compact) manifolds of dimension $n = 4, 5$ and, for these, the result is claimed also by R.Schoen [50] in a Lecture Notes Series (Montecatini Lecture Notes [50]), that is, in these Notes, R.Schoen proposes a proof of the positive mass conjecture for $n \leq 7$, non-conformally flat and this implies the general proof of the Yamabe conjecture, with the additional information that the solution is a minimum.

There is another proof of the Yamabe conjecture that we did establish in [11] and in [13] ([13] is a collaboration with H.Brezis , following the techniques of [14], [15]), in the cases left open by T.Aubin [6], but the solution derived in this way is not known to be a minimum.

Since this additional information has its merits, we want to analyze here the proof of the positive mass conjecture proposed by R.Schoen in these Montecatini Notes. We point here an important step in this proposed proof that is non-obvious to us. This step has been solved for the case $n = 3$ by T.Colding and W.Minicozzi II in [22]. The other cases, $n = 4, 5, 6, 7$, have not been solved in this monograph.

1. The proof of the positive mass conjecture as proposed by R.Schoen [50], graph behavior of the minimal surface and further estimates.

R.Schoen assumes in order to prove the positive mass conjecture in [50] that the conclusion of this conjecture does not hold at some point of M. R.Schoen then constructs a manifold $\hat{M}$ after sending this point to infinity that has one end at infinity carrying an asymptotically flat metric, which has a negative " mass" E. He then uses the negativity of the mass to build a sequence of minimal hyper-surfaces of codimension 1 Σ_σ having as a boundary an $(n-2)$-sphere defined in the $(x_1, ..., x_n)$ **global** variables near infinity by the formula $\Gamma_{\sigma,a} = \{x = (x', x_n), |x'| = \sigma, x_n = a\}$. For $n \leq 7$, this hyper-surface exists and is free of singularities. This is a non-obvious, earlier result which follows from the work of several mathematicians; we skip here the bit of history which would be required to trace the proof of this result; see eg [57], Chapter 7, Theorem 37.7, p221 for a proof for codimension 1 minimizing currents in R^n.

The mean-convexity of the boundary implies that this minimal hyper-surface is "jailed" near infinity between two hyper-planes $\{-a_0 \leq x_n \leq a_0\}$ and therefore, it is possible [51] to minimize over the value of $x_n = a$ the volume of the hyper-surfaces $\Gamma_{\sigma,a}$. One then find a bona-fide minimum value for $a = a(\sigma)$. In dimension 3, by the results of Meeks and Yau [36], [37], one knows that this hyper-surface is an embedded disk. In higher dimension $n \leq 7$, we know that it is free of singularities.

R.Schoen then takes a sequence of values of σ tending to infinity and, using results of Geometric Measure Theory, see H.Federer [28] and the much easier and pleasant to read Notes of Leon Simon [57], he derives that this sequence converges as σ tends to infinity to a codimension 1 absolute minimizer hyper-surface Σ_∞. This convergence follows from Chapter 7 (also section 37 of Chapter 7, which considers codimension 1 minimizing currents) of Leon Simon's Notes [57], Theorems 34.5 and 37.2 of Chapter 7 in particular, including graph convergence.

However, R.Schoen in [50] claims and uses a bit more: he claims that the sequence Σ_σ is made of hyper-surfaces that are graphs under the form $x_n = f_\sigma(x_1, ..., x_{n-1})$ with bounded gradient (these bounds are independent of σ), in these global variables outside a fixed compact set.

Once this graph form of the minimal surface is established, R.Schoen uses in [50] the minimal surface equation in graph form $x_n = f(x_1, ..., x_{n-1})$ and claims several estimates on the function f and its derivatives.

These estimates are essential for two reasons: the contradiction argument of [50] uses a dimension reduction induction where Σ_∞ replaces M. This requires very precise estimates on Σ_∞, its behavior and the behavior of the metric induced on it at infinity; this metric must be asymptotically flat etc. Also, in order to apply the Gauss-Bonnet formula on the open hyper-surface Σ_∞ and conclude the argument, one needs very precise decay estimates and convergence of various integrals on the non-compact manifold Σ_∞, see [50], [55].

The graph claim is therefore essential. Let us try to prove this claim through various available techniques[2]: :

---

[2]Professor Ernst Kuwert from the University of Freiburg, Germany, has informed us that a detailed proof of this step in the proof of the positive mass conjecture as outlined in [50] was included in his Diploma thesis. This thesis follows closely the various steps proposed in [50]. It has been published in 1990 at the University of Bonn and it was later referred to in one of his papers in 1991. The argument for the volume of balls goes as follows: one considers the complement V of a large cylinder $C(0, R_0)$ in the minimal surface Σ, where $C(0, R_0)$ contains the compact part of the manifold M. Let p be the projection on the plane H defined by the equation $x_n = 0$. Let

2. Some results from Geometric Measure Theory about graphs.

Looking at the results in Geometric Measure Theory, we find that there are Allard-type estimates (see W.K.Allard [1]) for the variation of the normal under specific conditions, see [57], Theorems 23.1 and 24.2 and conditions (22.4)and (24.1). These conditions depend, via the monotonicity formula, on the excess, see [1], [2], [28] and [57], and on the volume of the traces of euclidian balls on the hyper-surface (which can have high codimension). These estimates are certainly verified for very small balls, but very small balls (we cannot even estimate how small they are) are useless in order to conclude that Σ is a graph $x_n = f(x_1, ..., x_{n-1})$, with bounded gradient. Furthermore, these estimates are not available uniformly (no uniform upper-bound for the variation of the normal, no uniform upper-bound on the size of the balls where these bounds hold) on the sequence; they exist for each of the hyper-surface, but no uniformity follows from Leon Simon's Notes [57]; one can get this uniformity through the convergence process, but the bounds depend then on the limit hyper-surface, which is unknown. Therefore, a global L^∞ bound on the sequence of variations for the normal (and the size of the domains where this holds) or on the sequence of fundamental forms of the hyper-surfaces is not available through these techniques, let alone the fact that these hyper-surfaces are, outside a compact set, graphs $x_n = f_\sigma(x_1, ..., x_{n-1})$, with bounded gradient.

There are some more papers by R.Schoen and L.Simon about regularity of minimizers or stable extremizers of parametric functionals, see eg [51], [52] and [53].

[51] appears at first glance to be the ideal reference for use in the Riemannian (asymptotically flat) framework, but unfortunately the proof of Theorem 1 in [51] relies on (1.19), p 747 and that formula lacks the recording of an isometry. Once this isometry is recorded, (1.19) cannot be used in (4.5), p 765. The correct results are in [52]. In this paper, we can see that the variation of the normal is a bit less than Lipschitz, in line with the results of [1] and the results of [57]. The estimate also depends as above on an estimate on the excess and on an upper-bound on the volume of the traces of euclidian balls on the hyper-surface and our conclusion is that this is not general, for the reasons stated and described above.

[53] is another reference. It is a joint paper with F.J.Almgren and L.Simon, divided into two parts. The first one is a joint paper of R.Schoen and L.Simon. In this first part, the authors state Theorem 1.2, which claims a uniform Lipschitz control on the normal to a hyper-surface Σ of codimension 1 in R^n that minimizes an elliptic parametric

---

$K_\sigma = (H - C(0, R_0)) \cap \{x, |x| \leq \sigma\}$. First, observe that the projection from V to K is onto. This comes from either the "constancy theorem", see [57], p240, Theorem 41.1; or it comes from an algebraic topology argument: after generic perturbation, $p(V)$ defines a cycle in $(K_\sigma, \partial K_\sigma)$; its boundary value is not homologous to zero since one of its boundaries is the sphere of radius σ in K_σ and therefore, the cycle that it defines is the cycle of top dimension $(n-1)$ in $(K_\sigma, \partial K_\sigma)$. It cannot have "holes" then. So p is onto from V to K_σ. Consider then a ball of radius ρ around a point x_0 of V and consider $W = V - C(x_0, \rho)$, where $C(x_0, \rho)$ is the cylinder of radius ρ around x_0. The claim is that the area (for the conformally flat metric near infinity)of W is at least the area of $K_\sigma - B(p(x_0), \rho)$, up to a bounded constant. Indeed, since the projection p contracts distances (it is a projection) and is onto from W to $L_\sigma = K_\sigma - B(p(x_0), \rho)$, the **euclidian** area of W is at least the **euclidian** area of L_σ. We have to turn this into an estimate with the actual metric that we have. This metric is conformally euclidian, with the conformal factor $\phi(x) = 1 + A/|x|^{n-2} +$ higher order. Considering a point x in W and its projection ξ in K, we estimate the ration $\phi(x)/\phi(\xi)$. Because the last coordinate x_n is bounded, with uniform bounds, one can prove that this ratio is actually $1 + O(1/|\xi|^n)$. The function $1/|\xi|^n$ is integrable at infinity in H and therefore the area of W in the actual metric exceeds the area of L_σ in the actual metric, up to a negative bounded constant. On another hand, the area of Σ is less than the area of K_σ up to another constant: this is derived from the fact that Σ is minimal for all the spheres $S_h = \{(x', h), x' \in R^{n-1}\}$ for $-h_0 \leq h \leq h_0$, so that the area of Σ is less than the area of K_σ, up to a constant depending on R_0 only. We therefore find by difference that the area of the $\Sigma \cap B(x_0, \rho)$ is at most the area of the ball around $p(x_0)$ in H of radius ρ, which we can estimate directly, up to a bounded negative constant. The volume of the trace of the balls in Σ is therefore very close to the volume of the euclidian balls of the same dimension; the excess is estimated easily because the hyper-surface is jailed between two fixed hyper-planes. Σ is smooth so that θ of [57], p86, in Corollary 17.8 is 1 everywhere. All conditions are verified in order to use Allard regularity theorem under its first form, see [57], Theorem 23.1, p120. The proof is complete. Let us further observe that the estimate of the area of the trace of large balls on Σ stated in [55] and used in order to prove that $\int |A|^2 \lesssim \infty$, which we use below, see the section on the use of the "Stability formula and Simons inequality", actually does not hold if one does not know that the portion of the surface inside the ball is connected; one could have several "sheets" coming in and out of a given ball and building up more and more area for the portion of the surface inside the ball, even if it is absolutely minimizing. Thus, we do not really know before [50] that $\int |A|^2 \lesssim \infty$, even in the case $n = 3$, this according to our understanding. The proof of R.Schoen outlined in [50] is now, after the argument detailed above has been communicated to us-this has occurred only after an earlier version of this paper was posted amongst a few papers in our personal "website"-clearly complete according to our understanding and we can teach it to various students.

integrand, see [53]. The proof is not included in [53]. Several references are quoted for this result, for which a new proof is announced to appear in [52]. As stated above, such an estimate is not established in [52].

A (uniform) Lipschitz estimate on the normal is not the main focus of our present discussion: a stronger/different estimate is required in order to derive the graph behavior as stated by R.Schoen in [50], see eg (7.106) p253 of [22], where a decay in $1/|x|$ is established at infinity for the second fundamental form. This combines then with the fact that the metric is almost flat at infinity and the graph behavior as stated by R.Schoen in [50] follows.

Let us however, for the sake of completeness, digress for a short while from this main focus and discuss the references of [52] for the proof of Theorem 1.2.

The estimate of Theorem 1.2 is a celebrated result of E.De Giorgi, see p 260-265 in [25], who does indeed establish the Lipschitz statement of Theorem 1.2 for area-minmizing oriented boundaries of Cacciopoli subsets of R^n.

However, Theorem 1.2 of [53] claims in fact the same estimate under more general conditions. These more general conditions include the use of elliptic parametric functionals ([2], [28] and more). In the case of Riemannian manifolds of dimension $(n-1)$, we cannot assume that they are contained in R^n, but we can assume that they are contained(isometrically, using the Nash embedding theorem) in some euclidian R^N, with N large. We are therefore naturally led to results involving either elliptic parametric functionals which are different from the usual area functional or we are led to results involving higher codimension. Let us observe that R.Schoen and L.Simon did try in [51] to consider the case of elliptic parametric functionals in codimension 1; but this attempt did not produce a significant result as described above: the excess does not improve very well (ie following an exponential law) for parametric functionals as the size of the balls is decreased by eg "half". The recording of an isometry destroys the estimates, see [51] and [52].

The remaining frameworks involve higher codimension than 1 and also maybe parametric functionals. These two features produce the same type of results and they are therefore addressed together in the other references stated in [54] for the proof of Theorem 1.2. Fetching, we first find the book by H.Federer [28], Theorem 5.3.14 for the case of elliptic parametric integrands in arbitrary codimension: the "normal"(the codimension could be higher than 1) obeys a a $C^{1/2}$-estimate. This estimate holds again under conditions on the excess and on the volume of the traces of balls as above, with some differences and variants. However, even after considering these differences, we conclude as above that these conditions are not verified necessarily by Σ and we conclude as above that this estimate is useless for the strong and specific graph claim in [50].

Looking further, there is another reference which is a paper by F.J.Almgren [2]. In [2], for elliptic parametric integrands and in arbitrary codimension again, we find Theorem 7.5, which provides a Holder condition of order α on the normal, with $\alpha \leq 1/2$, see Lemma 7.4 and its requirements. Again, we find conditions on the excess and on the volume of the traces of balls which we do not know to be verified by Σ.

There is yet, as stated above, a third reference which is the work of W.K.Allard [1]. A Lipschitz estimate on the normal is nowhere to be found in [1]; a confusion with another statement, given in (8) of Lemma 8.12, p475 of [1], must be avoided. This statement says that the graph component of the hyper-surface over the hyperplane T approximating the varifold V (notations of [1]) satisfies a uniform Lipschitz condition, quite far and different from a uniform Lipschitz condition on the normal to V. All the other estimates by W.K.Allard in [1] on the normal to the hyper-surface are weaker than Lipschitz, in line with the estimates stated by L.Simon in his Lecture Notes [57], under more general assumptions than the assumption that the hyper-surface is minimal, see [57] as described in details above. They are also derived under assumptions as above on the excess and on the volume of traces of balls.

It is interesting to trace the reason why the "normal" to the hyper-surface is found, even in the case when this hyper-surface is minimal, to behave a bit less than a Lipschitz function. Looking at [57], Theorems 23.1 and 24.2 and conditions (22.4)and (24.1), we are led to think that, when the mean curvature H of the hyper-surface is zero, we can take in the estimates of this theorem $p = \infty$ and thereby derive a Lipschitz estimate on the normal. However, re-reading the proofs of these theorems, we find that they depend on a constant c computed (in part) in the proof of Lemma 19.5, p99 of [57]. Setting the mean curvature H to be zero in this lemma, we find that c is then, p 100 of [57], a fixed constant which tends to infinity with n and with the codimension k. β, in this lemma and after, is less than 1, in fact it is less than 1/2. Then, in the proof of Theorem 22.5 in [57], upon which the proofs of Theorems 23.1 and 24.2 depend, we find p119 the requirement $2c \leq (\beta/8)^2$ when $p = \infty$. It cannot be satisfied in full generality. This is one of the arguments where the assumption $p \lneq \infty$ is used, even when the mean curvature H of the hyper-surface Σ is zero.

The estimates and techniques described above, see also below, are or have become the classical estimates and techniques of Geometric Measure Theory. They are the techniques used to determine whether a minimal hyper-surface is a graph, over what domain and in what codimension. They all involve estimates on the volume of the traces of balls, estimates on the excess and on the multiplicity θ, see [57]. These are non-obvious estimates, especially for the estimates of the volumes of the traces of balls.

Beyond these estimates, the regularity of the minimizing hyper-surfaces belongs to the theory of Partial Differential Equations. The domain where these estimates hold is the domain over which these equations are verified. In the special case when these hyper-surfaces are determined to be graphs of codimension 1 over a given domain (given by the previous step, as above), these equations are elliptic second order scalar equations, for which scale-invariant $C^{k,\alpha}$ Schauder estimates can be derived. This is the natural extension of Geometric Measure Theory, see [22], pp252-255 for an example illustrating how Geometric Measure Theory and classical theory of Partial Differential equations combine in order to derive further significant estimates.

3. The Stability formula and Simons inequality.

Leaving aside Geometric Measure Theory and absolute minimizers, there is a road that is left: the use of the stability of these hyper-surfaces, the second variation formula [49],[54], [55], combined with Simons inequality, see eg [22]. There are several results in this direction and, for $n = 3$, T.Colding and W.Minicozzi II settle this question in [22].

a. The stability formula in R.Schoen and S.T.Yau's proof of the positive mass conjecture [55].

The use of the stability formula and the derivation of the estimate $\int_\Sigma |A|^2 \lneq \infty$ by R.Schoen and S.T.Yau in [55],-let us observe here, about this three dimensional proof, that the estimate of the area of the trace of large balls on Σ stated in [55] and used in order to prove that $\int |A|^2 \lneq \infty$ does not hold if one does not know that the portion of the surface inside the ball is connected; one could have several "sheets" coming in and out of a given ball and building up more and more area for the portion of the surface inside the ball, even if it is absolutely minimizing. Thus, if the graph claim as in [50] is not established, we do not really know that $\int |A|^2 \lneq \infty$ in the case $n = 3$, this according to our understanding-with A being the second fundamental form of Σ, is specific of the dimension 3 as some simple computations, carried when trying to generalize the argument of [55] for the estimate on $\int_\Sigma |A|^2 \lneq \infty$ to a higher dimension than 3, clearly demonstrate. The estimate in higher dimensions $n \geq 4$ (following the same arguments than [55], these arguments are not sound as noted above)reads:

$$\int_{B_{r_0}(0)\cap\Sigma} |A|^2 \leq C r_0^{n-3}$$

for r_0 large. $B_{r_0}(0)$ is the ball around zero of radius r_0 in the asymptotically flat ambient manifold M^n.

We can try then to follow the argument of T.Colding and W.Minicozzi II [22] for the proof of the positive mass conjecture in dimension 3. Up to some additional details, such as taking care of the constants in the Simons inequality, see eg [22], p66, formula (2.1), the argument relies on H.I.Choi and R.Schoen's estimate on the second fundamental form A of a minimal surface in [20]. [20] provides an L^∞ bound on the second fundamental form of a codimension 1-hypersurface (the authors in [20] write the estimate for surfaces. but the result generalizes)and this bound depends on an priori bound on the L^2-norm of the second fundamental form on a ball. Let us try to make this work here for $n \geq 4$: we know that the ambient manifold M is asymptotically flat near infinity, with a fast decay for the curvature. This allows to estimate the radius of injectivity at a point x_0 near the non-compact end of M to be of the same order than $|x_0|$ (it cannot be more, there is the compact part of M to take into account). Considering a suitably small fixed constant $c \gneq 0$, we would need, by scale-invariance and after copying the proof of [20] to the higher dimensions, an estimate telling us that

$$\frac{\int_{B_{c|x_0|}(x_0)\cap\Sigma} |A|^2}{(c|x_0|)^{n-3}} = o(1)$$

$o(1)$ means that this quantity is suitably small. This estimate holds for $n = 3$ when $|x_0|$ is large, with a fixed small constant c. This is due to the fact that $\int_\Sigma |A|^2 \lneq \infty$.

However, in higher dimensions, we only know, see above, that $\int_{B_{r_0}(0)\cap\Sigma} |A|^2 \leq Cr_0^{n-3}$. This comes very close, but it is not enough. Would we have the different estimate $\frac{\int_{B_{c|x_0|}(x_0)\cap\Sigma} |A|^2}{(c|x_0|)^{n-3}}$ small enough, then we could use the fact that the manifold M is asymptotically flat, Simons inequality as in [22] and derive that $|x||A(x)|$ is bounded above. This would then imply that the hyper-surface Σ is a graph $x_n = f(x')$ as stated by R.Schoen in [50], with gradient tending to zero at infinity. Using then the minimal surface equation in graph form and Schauder $C^{2,\alpha}$ estimates on an exterior domain in the variable x', we would find, as in [22], p254, that $|x|^2|A(x)|$ is bounded above, so that $\int_\Sigma |A|^2$ would be convergent for $n = 3, 4$. The positive mass conjecture would follow for these two dimensions, using the induction argument of R.Schoen in [50] and the Yamabe conjecture, with the additional knowledge when compared to the results derived in [11], [13], that there is a solution which is an absolute minimum for the Yamabe functional, see eg [6], would also follow in these two dimensions. The dimensions $5, 6, 7$ for the positive mass conjecture would require further arguments, see below.

b. Theorem 3 of "Curvature Estimates", Acta Mathematica 1975.

Trying to use other arguments, we find, for the general case $3 \leq n \leq 5$, a paper by R.Schoen, L.Simon and S.T Yau [54]. Estimate (2.9) of [54] reads $\frac{\int_{B_{c|x_0|}(x_0)\cap\Sigma} |A|^2}{(c|x_0|)^{n-3}} \leq \beta_1$, where β_1 is a suitable positive constant. This is a bit short of what is needed. On the other hand, Theorem 3 of [54] gives a local estimate on the second fundamental form, but there is an assumption on the curvature of the ambient manifold: it is non-positive. Our manifold does not have non-positive curvature; but it is flat at infinity, with a fast decay for the curvature. We are therefore led to re-read the proof of Theorem 3 in [54] and see if we can adapt it here. The task is not completely straightforward as we have to check the availability of a Sobolev inequality on the domain where we wish to use it. Such a domain can be traced back to be a ball in Σ of radius $c|x_0|$ around a point x_0 of Σ near infinity in M. As noted above, the radius of injectivity of M near $|x_0|$ is of the order of $|x_0|$. We check in [34], a work by D.Hoffman and J.Spruck, conditions (2.2), (2.3), (2.4) in Theorem 2.1 and we find that they are indeed verified. We therefore can try to extend Theorem 3 of [54] to our asymptotically flat manifolds. We carefully check the proof and we find that it relies on a form of the mean value inequality for which the authors of [54] appeal to Theorem 5.3.1 in C.B.Morrey [39]. This is p137 of [39] and its proof relies on the proof of Lemma 5.3.3, p135 of [39]. Of special interest are the conditions on d in the statement of this lemma, which the coefficient of the function W which we want to estimate in Lemma 5.3.3 of [39]. The condition on d is found in 5.1, formula (5.1.3), p126 of [39]. It reads:

$$\int_{B_{x_0}(r)\cap\Sigma} |d|^{\nu/2} \leq (C_0 r^{\mu_1})^{\nu/2}$$

The balls $B_{x_0}(r)$ can here be indifferently taken to be euclidian balls or intrinsic balls of Σ; it does not matter, the manifold is asymptotically flat, with a fast decaying curvature. ν is the dimension $(n-1)$ of Σ ($(n-1)$ is n in [54]).

In our case, $|d|$ can be recognized to be lower-bounded by $|A|^2$, A the second fundamental form of Σ, see (2.11) of [54]. The above condition therefore re-reads then $\int_{B_{x_0}(r)\cap\Sigma} |A|^{n-1} \leq (C_0 r^{\mu_1})^{(n-1)/2}$. Estimate (2.9) of [54] tells us that $\int_{B_{x_0}(r)\cap\Sigma} |A|^{n-1} \leq C$ and this scale-invariant cannot be improved with the use of larger ps, when available in (2.9) of [54]. This is the inequality (5.1.3) in our framework.

We now come to the argument of Lemma 5.3.3 of C.B.Morrey [39]. We find that we need for the proof of this Lemma the inequality $2CZ_2\tau^2 \leq 1$, notations of Lemma 5.3.3 of [39], $\tau \geq 1$ and Z_2 are constants. This is required in order to derive (5.3.15) of [39] from (5.3.12) and (5.3.14), that is in order to absorb in (5.3.12) the term $Z_2\tau^2 \int_G P^2 w_L^2$ of the right hand side into the left hand side after use of the Holder inequality and the Sobolev inequality. The condition is absolutely necessary and it requires therefore a smallness assumption on $\int_{B_{x_0}(a)\cap\Sigma} |A|^{n-1}$, which we do not have.

One could try to use further estimates by R.Schoen, L.Simon and S.T.Yau in [54]and modify the argument of C.B.Morrey in [39] to adjust for the present situation. For example, an estimate of the type $\int_{B_{x_0}(r)\cap\Sigma}|A|^p \leq Cr^{(n-1-p)}$ is established in [54] for $0 \lneqq p \lneqq 4+\sqrt{8/(n-1)}$. For $n=3,4,5,6$, this range exceeds the closed interval $[0,(n-1)]$ and therefore, the estimate on $\int |A|^p$ on balls of radius r goes beyond the exponent $p=(n-1)$. Inserting this in (5.3.14) of [39], p136, we use Holder with a higher exponent on P^2, we find a smaller exponent than the critical exponent on w_L; we use Holder again and the Sobolev inequality. The conclusion is the same, as expected: μ_1 is zero and the argument of [39], Theorem 5.3.1 does not apply without a smallness assumption on $\int_{B_{x_0}(a)\cap\Sigma}|A|^{n-1}$, which is almost equivalent to the smallness assumption in H.I.Choi and R.Schoen's work [20], see also above our observations about possible extensions.

The proof of Theorem 3 in [54] is not complete, according to us and the extension to asymptotically flat manifolds which we have tried to complete cannot be performed as such. Some more work is needed.

Repeating the outline for a proof of the positive mass conjecture in the non conformally flat case, as established by R.Schoen and S.T.Yau [55] for $n=3$-with the help of [22] to make this argument much more precise and rigorous in all details- and as proposed by R.Schoen [50] for the other dimensions $n=4,5,6,7$, assuming that such a result is established, then the estimate $|x||A(x)| \leq C$ would follow. This would imply the graph behavior $x_n=f(x')$, $|\nabla f(x')|$ small near infinity. This in turn would imply, using the minimal surface equation in graph form and $C^{2,\alpha}$-Schauder estimates that $|x|^2|A(x)|$ is bounded. This estimate would provide a decay in $1/|x'|$ for $|\nabla f(x')|$ and, from there, the other estimates claimed in [50] would follow after Kelvin transform and iterated use of the Schauder estimates (and a second reverse Kelvin transform to apply $C^{2,\alpha}$-estimates on exterior domains as in[22]). Then, a decay on $|A(x)|$ stronger than $1/|x|^2$ could be derived and this would imply that $\int_\Sigma |A|^2 \lneqq \infty$. The proof of the positive mass conjecture and the proof of the Yamabe conjecture in all the cases left open by T.Aubin [6] would then follow, with the additional information that there is a solution for the Yamabe problem that it is an absolute minimum for the corresponding functional[6]; this for $n=3,4,5,6$ to the least.

c.An additional reference.

There is yet an additional reference that we should add to our present discussion about the use of the stability formula: when $n=3$, there is a paper by R.Schoen [49] in Annals of Mathematical Studies 103, with a scale invariant L^∞-estimate on the second fundamental form of a stable orientable minimal surface. The proof in [49] uses a (reverse) Bishop type estimate in Theorem 3 and Corollary 4; this estimate is known to hold under Ricci curvature lower-bounds that are tantamount to the conclusion of the claims. T.Colding and W.Minicozzi II [22] have now established these results. R.Schoen's Theorem 3 and Corollary 4 of [49] hold true, but the proof today in [22] is different.

4. Conclusion.

There is no doubt that this is a very interesting detail[3] to fill in with a rigorous proof and that the tools available in the literature (and cited above) come close to such a proof; but this is not complete. More "details" would be very helpful in order to feel "comfortable" with the proof of the positive mass conjecture in these dimensions, the proof of the existence of an absolute minimum for the Yamabe variational problem in these dimensions as well (knowing that the conjecture is settled by the arguments of [11] and [13] as well, but without this minimum feature) and an additional thorough knowledge of the behavior of stable and absolutely minimizing hyper-surfaces near infinity.

[50] contains an additional section where the behavior of sequences of solutions to almost critical exponent equations is analyzed. There, a combination of Green's function at various poles appears. This direction has also been fruitful in the study of singular and regular solutions to limiting exponent equations [12], [45], [48]. This multiple points blow-up, **with the additional** use of the Green's functions at various points appeared within the framework of regular solutions to critical exponent equations in 1985 in our work with J.M.Coron [14], [15]. [15] is also the reference where the solution to the Kazdan-Warner problem on S^3 has been outlined in a detailed manner. The critical

[3] see our previous footnote for the solution of this "detail".

points at infinity are defined and recognized in [15] for the scalar curvature problem, their topological contribution is analyzed in details (with the support of analytical and dynamical estimates, see [14], [15]) and the Euler-Poincare characteristic argument is provided, leading to the existence of solutions, under the appropriate topological condition, see [15].

Coming back to regular and singular solutions of Yamabe-type problems, these techniques, originating and expanded through [14], [15], [45], [48], [12] and more (this is not an exhausting list. There are several papers following [14], [15] and [48] that we are not citing here. [48] is the main reference in these papers, but [45] is the first paper where the relationship between the regular and singular solutions of the Yamabe equation appears, see [12] in order to understand this link) have created a very fruitful tool. Since then, regular and singular solutions have developed in parallel and that is a very good feature in the field.

Multiplying the number of blow-up points and trying to track the behavior of the limiting critical configurations, as this number goes to infinity, is another related direction of research, that has been outlined in [12], [14], [15]. It is still open for Yamabe-type problems. On another hand, a definite progress in the understanding (as the number of blow-up points or singularities goes to infinity) has been completed in other problems having some similarities, see the work of E.Sandier and S.Serfaty [46] on Ginzburg-Landau equations.

Yamabe-changing sign equations [17] is an independent area of research which belongs more to Nonlinear Analysis than Geometry, but is part of the same general area of research. The associated variational problem at infinity for these equations is different from the problem at infinity for (positive) solutions of Yamabe-type problems. This variational problem at infinity has not been yet understood, see [17], the Introduction for a description of several features of this problem.

3. One of the concluding arguments for the proof of the three dimensional Poincare conjecture: the curve shortening flow on curves [43], [38].

G.Perelman [43], [44] and more has proposed a method in order to solve the three dimensional Poincare conjecture. His papers are sometimes sketchy and several mathematicians gathered to try to write expanded details based on his ideas.

The proofs that came out of this tremendous effort are extremely long and they are hard to check. We will discuss-and this is the only part that we have read in details-here solely the conclusive arguments in the proof of this conjecture for the vanishing of π_3 of the manifold (these arguments are more involved than the arguments for the vanishing of the second homotopy group π_2. The latter arguments can therefore be considered contained in the arguments for the vanishing of π_3).

Two arguments have been proposed, one formally by G.Perelman [43], written in much details by J.Morgan and G.Tian [38]; the other one has been suggested to T.Colding and W.Minicozzi II [22], [23] by G.Perelman informally. This argument has been established by these two authors in [22], [23].

We show here that the first line of proof proposed by G.Perelman in [43] and carried out by J.Morgan and G.Tian in [38] has serious gaps:

Both proofs rely on studying a non trivial map from $S^2 \times S^1$ into M (a map of non-zero degree). In the line of proof proposed in [43], this is viewed as a non-zero class $[\sigma]$ in $\pi_2(\Lambda(M))$, $\Lambda(M)$ being the loop space of M. M is the three dimensional manifold that has now no π_2 after the use of the Ricci flow [21], [31], [38], [44] and more.

G.Perelman [43] associates to such an element $[\sigma]$ a min-max value, derived after computing the area of minimal parametrized disks that the curves of a representative c of $[\sigma]$ bound.

Maximizing the area for all curves of c, then minimizing over all c of $[\sigma]$ provides a critical value cv.

As the metric evolves through the Ricci flow, this critical value is a lower semi-continuous function of t, $cv(t)$. The claim, derived through the use of the Gauss-Bonnet formula [26] for minimal disks-after the use also of the curve-shortening flow-is that $cv(t)$ becomes negative in finite time: a contradiction.

The Gauss-Bonnet theorem on a disk has two terms: one term is an inside contribution; it is a surface term, equal to $\int Kda$; K is the gaussian curvature of the disk. It is provided by the evolution of the metric, see Theorem 11.1 [32], also, for a closed surface, see Claim 18.12 of [38], pp 426-427. The second one is a boundary term, $\int k_{geo}ds$, where k_{geo} is the geodesic curvature of the boundary curve. It has to be forced into the formula and this is obtained through the use of the so-called curve shortening flow H on curves, see Lemma 19.2, pp 438-439 of[38].

This curve-shortening flow has several problems, unwanted blow-ups [3], [5]; these problems are clearly identified and studied in the literature [3], [5], [30] for the case where no singularities arise etc. This flow exists under another form in Contact Form Geometry (see [10], the flow with $\eta = b$). In Contact Form Geometry, this flow is clearly undesirable as the small loops that might form on the curves under deformation yield through the lift that occurs in Contact Form Geometry (eg from a curve on a surface Σ to a curve on its unit sphere cotangent bundle, see eg [7], the introduction)large loops that cannot be dealt with.

However, it has been suggested by Calabi and carried out by Altschuler-Grayson [4] that lifting this flow from Σ to $\Sigma \times R$, or as is carried out by J.Morgan and G.Tian [38], following the suggestions of G.Perelman [43], Σ to $\Sigma \times S^1$ with a small length ϵ for the curves along the S^1-component, allows to bypass these unwanted blow-ups of the curve shortening flow.

This flow is in this way-it takes some work to see this-deemed to have a weak-limit continuation through its blow-ups and, in addition, after some "averaging", this designs a continuous process in the space of immersed curves of the three dimensional manifold M.

This is what G.Perelman asserted in [43] and this was claimed to hold with much details carried on in the work by J.Morgan and G.Tian [38]. The general idea follows the estimates of Altschuler-Grayson in [4], but the claims of [38] and [43] go beyond [4] which is only concerned with planar curves. In [4], the computations are carried for the **algebraic** curvature k; they are not carried for k^2, with $k = |k|$ as in [38], [43], for three dimensional curves (for which an algebraic curvature k is not defined).

We prove here that these claims are wrong and that there might be a serious obstruction to reach such a conclusion.

The argument by G.Perelman [43] is sketchy and leaves several details unchecked. J.Morgan and G.Tian [38] did carry these computations into much detail and made the computations of G.Perelman, rather his sketches, very precise. They reached the following inequality in Lemma 19.9 page 444 of their monograph [38]. k^2 is the square of the norm of $\frac{dT}{ds}$, T is the unit tangent vector to the curve, s is arclength; the metric is evolving along the Ricci flow and the boundary curve is evolving through the curve shortening flow H, see [38], p 437:

$$\frac{\partial k^2}{\partial t} \le (k^2)\text{"} - 2((\nabla_S H)^\perp, (\nabla_S H)^\perp))_{c^*g} + 2k^4 + \hat{C}k^2 (*)$$

$(*)$ above computes precisely the missing terms in formula (3) of G.Perelman [43]. G.Perelman derives then in [43] formula (4). This involves an unjustified division by k.

J.Morgan and G.Tian [38] also divide by k and derive, see Claim 19.8 p 443 (k is non-negative, but k can be zero):

$$\frac{\partial k}{\partial t} \le k\text{"} + k^3 + C_1 k$$

This gives rise to an exponential law on $\int kds$(k^3 disappears after integration because the metric is evolved by the Ricci flow so that the arclength parameter s is time-dependent). Therefore, if a closed curve is a geodesic at time t_0, it remains so for $t \ge t_0$ since H is zero on this curve and the above inequality implies a (dominated) exponential law on $\int kds$.

This has of course easy counterexamples.

The exponential law is stated explicitly in Morgan-Tian [38], Lemma 19.9, p 444.

This inequality is, in turn, used on time intervals jailing the times where the projections of the ramp (see [38], p 445) solutions on M do not converge. The only knowledge on these time intervals is that their total measure is as small as we wish, the smallness is related to how large are the bounds on the complement intervals; we must take a fixed size of smallness for the total measure of these intervals. However, this information, combined with this unjustified wrong law (would $\int kds$ not be a priori bounded, we would run into the difficulty of $0 \times \infty$!) allows to conclude , see [38], Claim 19.23 and its proof which refers to Lemma 19.9 explicitly, p 453-454. Claim 19.23 is in turn used in the proof of Lemma 19.25, p461, for the conclusion of the proof of this Lemma. It is a fundamental estimate.

Correcting this may take some work: assuming that we have immersed curves along the curve shortening flow having $k(x,t)$ positive for each x and for t in some time interval, then Corollary 19.10 holds for them. But the exponential inequality on Θ can be extended to curves in their closure in the C^2-sense. It follows that geodesics cannot be close to them in the C^2-sense. The inequality becomes very unstable or the homotopy type of the compact set of curves is not general.

A clear mistake is happening here with the unjustified division by k and this cannot be corrected easily. There are serious obstructions to extend weakly the curve-shortening flow to all curves.

We had reached such a conclusion in Contact Form Geometry (for the flow defined with $\eta = b$, [10]) and we had "discarded" this flow and built another flow, see [10].

4. Floer Homologies via gauge Theory and Deformations of the Seiberg-Witten equations.

The Seiberg-Witten equations have been extensively studied and explored over the last twenty years, with a wide variety of applications, from the definition of invariants for three and four dimensional manifolds eg [35]) to the proof of the Weinstein conjecture in dimension 3 by C.H.Taubes [59].

We claim that the construction by P.Kronheimer and T.Mrowka [35] for three flavors of Floer homology using these equations must be an S^1-equivariant theory, involving "point to circle" Morse relations [9]; but because the compactness of the instantons of the theory of P.Kronheimer and T.Mrowka [35] occurs only after gauge re-normalization in their construction, see Theorem 5.1.1, p100 of [35] and because the results in [35] do not provide estimates on this re-normalization from $-\infty$ to ∞ once a prescribed gauge is given at $-\infty$, the "point to circle" Morse relations involving circles of instantons [35] originating at the same irreducible SW-solution are not understood; some other phenomena may occur which would lead to define other critical subsets of a circle of SW-solutions besides the points of the circle and the whole circle (these are the "top" and "bottom" critical points on a circle; some other subsets of a circle could be "critical" if asymptotes, due to the gauge re-normalizations, develop along instantons as the circle parameter varies): an S^1-equivariant theory is not available then, within the current knowledge.

The proof of the three dimensional Weinstein conjecture by C.H.Taubes [59] depends in a crucial way on the construction of [35], on $\hat{HM}$ in particular, see [35]. Since this construction is not complete, the proof by C.H.Taubes [59] is not complete.

To make things easier, we recall (it is a first step, to be refined later) that the gauge group in the Seiberg-Witten equations is $C^\infty(M, S^1)$ or some Sobolev form of this space. Assuming that M is simply connected, this space can be topologically identified with S^1 (action on the A-part of the solution is $-u^{-1}du$, action on the Ψ-part is $u\Psi$, $u \in S^1$, see [59] for example).

Irreducible Seiberg-Witten solutions provide then circles of solutions, whereas reducible solutions are fixed under this circle action. This reads through the Ψ-component of the (A, Ψ) description of a Seiberg-Witten solution; A is a connection over a "difference" line bundle over M^3 and Ψ is a section of the associated bundle $S \underset{U(2)}{\times} C^2$ to the $U(2)$-bundle S over M^3 derived after lift of the frame bundle from an $O(3)$-bundle to an $SO(3)$-bundle (because M^3 is orientable) to a $Spin(3)$-bundle (M^3 is a three dimensional orientable manifold) and then to a $Spin(3) \underset{Z_2}{\times} U(1) \cong U(2)$ bundle S.

Ψ is zero at a reducible Seiberg-Witten solution. It is a fixed point for the S^1-action. A small neighborhood of zero in the space of the sections of S yields as a boundary a small sphere in an appropriate Sobolev space with a free action of S^1. The quotient is a copy of PC^∞. It is used to provide, eg for $M = S^3$, the appropriate framework in order to compute these three homologies $\hat{HM}$, $\bar{HM}$, $\check{HM}$.

The "invariants" are defined, we are developing this in the case of a single reducible Seiberg-Witten solution, by considering the "flow-lines" of the Seiberg-Witten functional, viewed as moduli spaces over cylinders (that break). This happens on the "blown-up space ", see [35], blown-up space of the space of couples (A, Ψ). The blow-up occurs mainly in the Ψ-space, but involves in fact via the Dirac operator (see below) associated to the connection A the two components. This blow-up is completed in order to resolve the fixed point set of the S^1-action.

Roughly speaking, this develops Morse Theory for the "compact" (there are some gauge re-normalizations to be completed, they actually play a central role in the theory, see below) Seiberg-Witten functional under the form of moduli spaces (diffusion equations). This Morse Theory is available [35] on the "blown-up space", a manifold that has as boundary (this is an approximation)a small sphere around the reducible Seiberg-Witten solution in the Ψ-space.

Near the small sphere "boundary" (this is occurring only in the Ψ-space, near the reducible SW-solution in particular), the variational problem is essentially understood using the eigenvalues and eigenfunctions of the Dirac operator D_{A_0} (an algebraic root for the Laplacian, see eg [40])in the Ψ-space (the reducible Seiberg-Witten solution reads $(A_0, 0)$) whereas in the space of connections A, there is (locally) an "almost" independent variational problem.

The positive eigenvalues/eigenfunctions of the Dirac operator provide the "stable reducible connections " C^s ([35], p411), whereas the negative eigenvalues/eigenfunctions of this Dirac operator provide the "unstable reducible connections" C^u. C^o is the space of irreducible solutions. All of this explained in [35], page 410 and after.

For a Morse function, the decomposition and the relative stratification of dominating and dominated stable and unstable manifolds has been understood in [16], Proposition 7.24 p 608 in particular and pp 604-624.

P.Kronheimer and T.Mrowka [35] extend this to the framework of moduli spaces in the Seiberg-Witten framework (a non-trivial matter); this framework has its own features and also they use diffusion equations (for non-monotone operators)/moduli spaces instead of pseudo-gradients. Within the "compact" (there is this gauge invariance that one has to take care of)framework, the pseudo-gradient approach and the diffusion approach yield similar result: the basic background is the same.

However, in the Seiberg-Witten framework, there are important differences as the linear operators are not monotone and their spectrum is not bounded above and below. The ellipticity of the operators, after modding out by the action of the gauge group for example on a slice, see [35], and the use of the implicit function theorem combined with the "usual" index theorems for Fredholm maps allow to derive the result.

1.Dynamical Aspects.

We will discuss later the analytic aspects of this problem. Let us discuss at this point the dynamical aspects, as if the instantons' equations in [35] were equations of a dynamical flow, of a "diffusion pseudo-gradient":

It is interesting to observe first that the domain with boundary used by P.Kronheimer and T.Mrowka [35] in order to define their invariants-Seiberg-Witten "diffusion pseudo-gradient" in space, flow-lines of the Dirac operator on the boundary, they "fit"-does not obey the isolating block condition introduced by C.C Conley and R.Easton [24] in order to define the Conley index.

Indeed, considering two irreducible SW-connections, one dominating the reducible connection, the other one dominated by this reducible connection, we find that a "pseudo-gradient" flow will be tangent somewhere along the small sphere boundary from inside. This tangency from inside violates the isolating-block condition of C.C Conley and R.Easton [24].

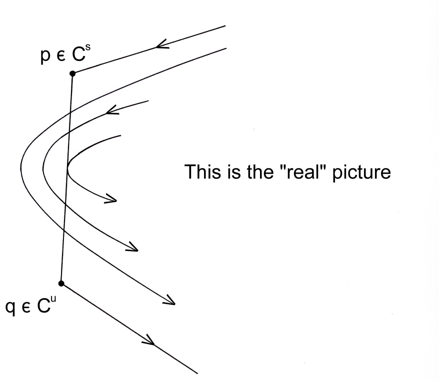
It might be that this is an instance where this fundamental condition is not needed. We will see that this is not true.

In fact, something is remarkable in P.Kronheimer and T.Mrowka's construction: no intersection operator ∂_o^s from a stable SW-reducible solution to an irreducible SW-solution is defined in their framework, see [35], p414.

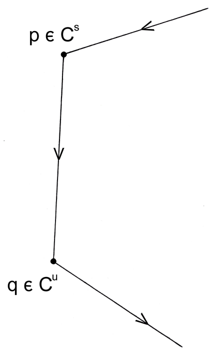
Does this mean that there is no such Morse relations with difference of indexes (spectral flows) equal to 1?

The answer to this question is negative: there are such Morse relations. However, in the framework of P.Kronheimer and T.Mrowka [35], they do not appear because this framework, understood in an extended way as the use of a "diffusion pseudo-gradient", imposes an algebraic formalism that violates transversality. When the index of the reducible SW-solution is given, then this can have a marginal impact on the invariants as the violation will induce dominations by the "stable reducibles" (using the terminology of [35]) of irreducible SW-solutions with the whole circle S^1 associated to each of them. Therefore, if one works modulo the action of S^1, this will not matter. However, when the index of the reducible SW-solution changes, since there are exit sets through its unstable manifold and since this unstable manifold changes rapidly its behavior over the change of index, some additional verifications are required (some non-obvious arguments-the arguments have to be completed **over** the bifurcations, not only before and after-, which we will not present here, lead to the conclusion that this "works").

Let us first explain why the algebraic formalism is not true to the dynamical phenomenon carried by a pseudo-gradient: any decreasing pseudo-gradient for the Seiberg-Witten equations can have flow-lines that "exit" the given boundary and then re-enter etc, yielding tangencies from inside:



After "algebraization", the picture for the flow-lines out of C^s becomes:



Flow-lines starting at $p \in C^s$ now have to go to a $q \in C^u$ in order to exit the boundary. There is no other exit set.

Then the loss of index at the boundary is 2 to the least. $q \in C^u$ has one more direction of index due to the fact that it is unstable; therefore between p and q, q with its entering (inside the domain where we are tracking flow-lines, out from the reducible) additional (half)-direction there is already a loss of index equal to 1. Any further domination would give a loss of index equal to 2 to the least, therefore no contribution to ∂_o^s.

This is not the real picture; it is the picture once the flow is **forced** into a non-transversal form where the stable manifold of an irreducible may develop a thick part that collapses at q, so that transversality does not hold, see the pictures above. One can see this over a process where the flow, with its tangencies-for which ∂_o^s might very well be non-zero-is forced into a flow for which all flow-lines from p exit only from qs. For transversality, differentiability (important for transversal intersections etc), the exit sets have to be open in the boundary, with tangencies etc: the dynamical picture is that flow-lines can come from an irreducible, go to the reducible SW-connection and from there, there can be unstable flow-lines going to a second dominated irreducible. Then, a subset of the set of flow-lines connecting the dominating and the dominated irreducible will exit through a small sphere around the reducible and also re-enter through the small sphere. The exit sets (into the exterior of the sphere, into the "blow-up" space where we are defining the invariants) must contain open subsets of the boundary for differentiability and this forces flow-lines outside of stable reducibles to exit not only through the unstable reducibles.

However, if the boundary is very, very small, the exiting flow-lines will be very close to those of the reducible SW-solution in the full space. This reducible SW-connection has an unstable manifold that is S^1-invariant. Some reasoning forces then that a stable reducible as in [35] can dominate an irreducible indeed, but it will be one that contains the full S^1-associated circle; that is irreducible SW solutions come as a full circle which can be thought as a top and as a bottom critical point. Here, the top critical point will be dominated and ∂_o^s is zero if we consider classes modulo S^1. Understood modulo S^1, with reducible SW-solutions that have a fixed relative Morse index or spectral flow, the construction of [35] appears sound from the dynamical point of view-despite the fact that it violates the isolating-block condition of C.C.Conley and R.Easton [24] as described above-if the sphere is taken very small and all is "stable" at the reducibles, that is their relative indexes (spectral flows) do not change.

As the parameter r of [59] changes, the reducible SW-solution undergoes bifurcations. One has to verify that the exit sets do not change "nature" (eg become not equivariant as the bifurcation occurs, given the fact that they are equivariant at the beginning of each bifurcation and after each bifurcation ends) over these bifurcations, so that the homology remains unchanged. This is not completed in [35], but it holds true nevertheless.

The conclusion of this sub-section is that an S^1 or gauge equivariant theory is needed; the dynamical aspects of the problem require such a construction. This is in agreement with the construction of [35], even though the phenomena are different when taken through our understanding, see the discussion above.

2.The Analytic Aspects.

Returning now to the analytic aspects, we identify in the Seiberg-Witten functional two terms. One is the Chern-Simons functional $\int_M \hat{a} \wedge d\hat{a}$, where the connection A reads as the sum of a reference connection A_E, see eg [59], and of $\hat{a}$

$$A = A_E + \hat{a}$$

The second main term reads $\int_M \Psi D_A \Psi$. D_A, see above, is the Dirac operator associated to A acting on sections of the associated bundle (to the bundle S, see above) $S \underset{U(2)}{\times} C^2$.

Both functionals suggest the use of the space $W^{1/2,2}$ for $\hat{a}$ and Ψ and they give rise to linear self-adjoint operators with spectrum unbounded from above and from below. This holds when we consider a section (or "slice" see [35]) to the gauge group. There is in the framework of the Seiberg-Witten equation a global slice to the gauge group action, see [35].

The gauge group becomes then $W^{3/2,2}(M, S^1)$, the gain of one derivative is due to the fact that an element u of the gauge group acts on A using the formula:

$$A - u^{-1}du$$

Other tools to use here are the fact that $W^{1/2,2}$ injects in L^3 in dimension 3, so that $|\Psi|^2$ is in $L^{3/2}$ and the fact that $W^{1,3/2}$ injects in $W^{1/2,2}$.

Some non-obvious work [18] shows that the space $W^{3/2,2}(M, S^1)$ has the topology of the circle S^1. This is related to the existence of a lifting $u = e^{i\phi}, \phi \in W^{3/2,2}(M, R)$ for $u \in W^{3/2,2}(M, S^1)$. The work of [18] goes much beyond the results required here as it addresses the existence of lifts for maps in $W^{1,p}(M, S^1)$. This has been investigated further in other papers, see eg (this is, by far, non-exhaustive)the work of Bourgain-Brezis-Mironescu [19].

Thus, the analytic aspects confirm that, up to deformation of the gauge group, we are considering an S^1-equivariant theory. The S^1 subgroup of the gauge group is defined by $u = e^{i\theta}$, with θ a constant; whereas its "complement" subgroup is defined by the equation $u = e^{i\theta}, \int_M \theta = 0$. This decomposition is used below, in the study of the "point to circle" Morse relation.

3."Point to circle" Morse relations and the Seiberg-Witten framework.

Let us consider a variational problem which is S^1-equivariant and compact (verifying the Palais-Smale condition). An S^1-equivariant theory for this variational problem is typically derived after modding out by the action of S^1 and therefore this variational problem considers each circle as a critical manifold, with a top and bottom critical point. A subset of a circle of critical points that would not be a point or the entire circle should not be critical then. Furthermore, a central tool in this theory is the notion of "point to circle" Morse relation [9]: along such a Morse relation, a critical point of (strict) index m, x_m, made of a point on a circle, can dominate a critical point made of a circle of (strict) index $(m-2)$, x_{m-2}. Typically, this will happen if there is no "intermediate" critical point (made of a point) of index $(m-1)$, ie no critical point x_{m-1} dominated by x_m and dominating x_{m-2}. Then, the "point to circle" Morse relation between x_m and the circle defined by x_{m-2} occurs as the result of the fact that the unstable manifold of x_m (with a base point on the corresponding circle) dominates, with an intersection number equal to 1 or with a non-zero intersection number, the stable manifold of $S^1 * x_{m-2}$, the circle associated to x_{m-2}, based at a given point $x_{m-2}(\tau_0)$. As τ_0 runs into S^1, the compactness properties of the variational problem and the non-existence of an intermediate x_{m-1} imply that the degree does not change. The "point to circle" Morse relation follows from this. This is a central tool in the S^1-equivariant Morse theory and it takes its roots in three features: the S^1-invariance and the compactness of the variational problem on one hand, the fact that the only critical points on a circle of critical points are the points of the circle and the entire circle.

This last feature is obvious in a finite dimensional compact framework, with an S^1-invariance. As the variational problem becomes infinite dimensional and the gauge group becomes non-compact, even though having the topology of S^1, this last feature becomes non-obvious and can even become outright wrong.

Considering now the Seiberg-Witten framework, we want to develop such a theory. In order to "resolve" the gauge invariance, we impose for example an evolution equation on the gauge components: we require that the connections belong eg to Coulomb slices or are restricted to stay in the global slice to the gauge group action see [35]. We then observe that it is difficult, after this restriction, within this theory to track the evolution of a gauge-component along a **sequence** of instantons (that might not converge after this gauge fixing to an instanton).

Indeed, the manifolds of instantons result directly from the use of implicit function theorem, after warranting transversality of a Fredholm map of a certain index (see Chapter IV, sections 13 and 14 of [35], for the infinite cylinders, also Chapter V, section 17, for the theory on finite cylinders and the use of the Atiyah-Patodi-Singer variant of the index theorem, again in [35]). The Fredholm and transversality aspects warrant manifold-behavior of the various sets of instantons and local compactness of these sets. However, a family of instantons (with a changing asymptotic behavior in our case, eg changing along the circle of a given irreducible SW-solution), might not converge to an instanton. Then, the Fredholm and transversality results of Chapter IV and V of [35] do not provide any control on the gauge re-normalizations of this family of instantons since no limit is **a priori** available.

Turning now to the equations themselves and considering one given instanton, there is no evolution equation because the linear operators are not monotone, on the contrary they have a spectrum that is unbounded from above and from below; nevertheless, control on the gauge-component is derived in [35] in the vicinity of an SW-solution, using non-degeneracy and Coulomb slices, see [35], Chapter IV, section 13, Proposition 13.6.1, p237 in particular, also Chapter 17 and 18, 18.4, in particular.

Compactness of instantons, without prescribing a slice, does hold within the Seiberg-Witten framework, see Theorem 5.1.1, p 100, Theorem 10.9.2, p169, Theorem 13.3.5, p225, Theorem 16.1.3 and Propositions 16.1.5 and 16.1.6, pp276-277, all in [35], but only after gauge re-normalization and not over the cylinder in its entirety: this compactness is established in [35] on compact subsets of the cylinder in the initial versions, it is later extended, see Proposition 13.6.1, p 237, to subsets of the cylinder that might contain one of $-\infty$ or ∞. Proposition 13.6.1 is stated for one instanton, but it can be extended to a sequence of instantons.

This falls short of the estimates that are required: What we are seeking is compactness for a continuous family of instantons, in a given slice, from a given base point of the dominating irreducible SW-solution along its associated circle (with a prescribed, not changing gauge, given by the slice) to a running base point along the circle associated to the dominated SW-solution (again the gauge is prescribed by the slice, transversally to the base point which is running along the circle).

Since we cannot assert that the gauge component is "controlled" (that would be in fact equivalent to full compactness within the global slice, without gauge re-normalization, at least as the base point of the dominated SW-connection runs along the associated circle), we cannot track a degree as the base point moves along the circle of the lower SW-solution R_{m-2} (the index $(m-2)$ defined above is now replaced by the spectral flow at the SW-solution R_{m-2}, whose value at this solution is assumed to be $(m-2)$): the gauge component could "blow-up" whereas the "point to circle" Morse relation is not achieved, ie blow-up occurs at certain times τ along the circle, impeding the degree argument to proceed. There would then be other critical points or rest points along a circle besides the points of the circle and the full circle as some other phenomena, of variational nature, with stable and unstable manifolds, do occur.

If these additional critical points did induce differences of topology in the level sets, then they would be "essential". The critical sets of the Seiberg-Witten functional [35], [59] are non-compact. Assuming that the violations of the "point to circle" [9] Morse relations are essential, then the action of compactifying these critical sets through appropriate perturbations would not help much as the problems would then re-appear, hidden somewhere in the theory with the technical construction of the compactification added to it. If a family of instantons does not converge to an instanton, then there might very well be convergence when we mod out by the gauge group. But, some relevant phenomenon, with actual variational implications, might be happening there and this requires full attention and study before being discarded.

In conclusion, the existence and the compactness of these "point to circle" Morse relations (and the exclusion of other ones that occur along only part of the circle) is not established within the theory of [35]. It follows that the proof of the Weinstein conjecture by C.H.Taubes in dimension 3 [59] is not complete according to us. Some additional work is required: there is the need to control the gauge re-normalizations along a sequence of instantons involved in a "point to circle" Morse relation[9].

5.The fifth gap: pseudo-holomorphic curves and contact forms.

The fifth gap that we want to recall here occurs in the field of pseudo-holomorphic curves, when trying to solve the problems of periodic orbits for Reeb vector-fields "in the large". We had pointed, with our techniques and our language, some conceptual mistakes of this field in at least two papers[8],[9].

The independent work of JoAnn Nelson, see eg her PhD thesis [40], independent and **within** the field of pseudo-holomorphic curves, has confirmed the existence of compactness issues related to what we have called "point to circle" Morse relations in Contact Form Geometry. We have described these above, see also [9].

Conclusion.

As gaps and missing details multiply, the understanding decreases and the process of original creation finds itself with increased difficulties of expression, also of doubt as to the open areas and the problems that have or haven't been solved. The vision becomes "blurred".

In addition, the modern techniques in Mathematics are sometimes very sophisticated; they require a large amount of knowledge that is not readily available, even to the experienced researcher.

As a result of all this process, direct knowledge of the results and knowledge of what is true decrease and they are replaced by the much weaker second-hand knowledge and also unfortunately by error. The process of creation becomes also less genuine, second hand creation or creation based on approximative (or wrong) knowledge become more important, adding to the problems.

It is hard to see how this can be changed; we all probably have to go through some kind of criticism of our work and we all have to start to change.

All of these defects are of course companions to the increased speed and increased "movement". Left unchecked, these features can change Mathematics or certain areas of Mathematics to the least into a speculative field, a process already under way in some directions.

Some corrections would be welcome.

References

1. W.K.Allard, *On the first variation of a Varifold*, Annals of Maths, Second Series **95, no 3417-491** (1972).
2. F.J.Almgren, *Existence and Regularity Almost Everywhere of Solutions to Elliptic Variational Problems*, Annals of Mathematics, Second Series **87, No. 2** (1968321-391).
3. S.Altschuler, *Singularities of the curve shrinking flow for space curves*, J. Differential Geometry **34** (1991), 491514.
4. S.Altschuler and M.Grayson, *Shortening space curves and flow through singularities.*, J. Differential Geom. **35** (1992), 283-298.
5. S.Angenent, *On the formation of singularities in the curve shortening flow*, J. Differential Geom **33 no 3** (1991), 601633.
6. T.Aubin, *The Scalar Curvature*, Differential Geometry and Relativity , Mathematical Phys. and Appl. Math. (Cahen and Flato eds) Reidel, Dordrecht **3** (1976), 5-18.
7. A.Bahri, *Pseudo-Orbits of Contact Forms* (1988), Pitman Research Notes in Mathematics Series No. 173, Scientific and Technical, London.
8. A.Bahri, *Attaching Maps in the standard geodesics problem on* S^2, Discrete and Continuous Dynamical Systems 30, Number 2 (2011).
9. A.Bahri, *Topological Remarks-Critical Points at Infinity and String Theory*, Advanced Non-linear Studies **9** (2009), 439-454.
10. A.Bahri, *Classical and Quantic periodic motions of multiply polarized spin-manifolds.* (1998), Pitman Research Notes in Mathematics Series No. 378, Longman and Addison - Wesley, London and Reading, MA.
11. A.Bahri, *Another proof of the Yamabe conjecture for locally conformally flat manifolds*, Nonlinear Analysis: Theory, Methods Applications **20** (1993), no. 10, 1261-1278.
12. A.Bahri, *Recent progress in conformal geometry (with an emphasis on the use of configuration spaces) in "Lectures on Partial Differential Equations Proceedings in Honor of Louis Nirenbergs 75th Birthday"*, New Studies in Advanced Mathematics, 2003.
13. A.Bahri and H.Brezis, *Non-linear Elliptic Equations on Riemannian Manifolds* **20** (1996), PNLDE, Topics in Geometry, In Memory of Joseph D'Atri, 1-100.
14. A.Bahri and J.M.Coron, *Une theorie des points critiques a l'infini pour l'equation de Yamabe et le probleme de Kazdan-Warner*, C. R. Acad. Sci. Paris **300 Serie I, 15** (1985), 513-516.
15. A.Bahri and J.M.Coron, *Vers une Theorie des Points Critiques a l'Infini*, Seminaire Bony-Sjostrand-Meyer, Expose (VIII) (1984).
16. A.Bahri and P.H. Rabinowitz, *Periodic solutions of hamiltonian systems of 3-body type*, Annales de l'institut Henri Poincar (C) Analyse non linaire **8, no 6** (1991), 561-649.
17. A.Bahri and Y.Xu, *Recent Progress in Conformal geometry*, Imperial College Press Advanced Texts in Mathematics, London, 2007.
18. F. Bethuel and X. Zheng, *Density of smooth functions between two manifolds in Sobolev spaces*, J. Funct. Anal. **80** (1988), 6075.
19. J. Bourgain, H. Brezis and P. Mironescu, *Optimal Control and Partial Differential Equations, A Volume in Honour of A. Bensoussans 60th Birthday*, IOS Press, 2001, pp. 439455.

20. H.I.Choi and R.Schoen, *The space of minimal embeddings of a surface into a three-dimensional manifold of positive Ricci curvature*, Invent. Math **81 no 3** (1985), 387394.

21. B.Chow and D.Knopf, *The Ricci flow: an introduction*, vol. 110, Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2004.

22. T.Colding and W.P.Minicozzi II, *A Course in Minimal Surfaces*, vol. 121, Graduate Studies in Mathematics, AMS, 2011.

23. T.Colding and W.P.Minicozzi II, *Width and finite extinction time of Ricci flow*, Geometry and Topology **12** (2008), 2737-2786.

24. C. Conley and R. Easton, *Isolated invariant sets and isolating blocks*, Trans. Amer. Math. Soc. **158** (1971), 35-61.

25. E.De Giorgi, *Selected Papers*, L.Ambrosio, M.Miranda, G.Dal Maso, S.Spagnolo editors, Springer, 2006.

26. M. P. Do Carmo, *Riemannian Geometry*, Birkhauser, Boston, 1993.

27. S.K.Donaldson, *An application of gauge theory to the topology of 4-manifolds*, Jour.Diff.Geom **18** (1983), 269-316.

28. H.Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin-Heidelberg-New-York-yr 1969.

29. D.Freed and K.Uhlenbeck, *Instantons and Four Manifolds*, vol. 1, MSRI Publications, Springer-Verlag, New-York-Berlin-Heidelberg-Tokyo, 1984.

30. M.Gage and R.S.Hamilton, *The heat equation shrinking convex plane curves*, Journal of Differential Geometry **23 no 1** (1986), 69-96.

31. R.S. Hamilton, *Three-manifolds with positive Ricci curvature*, J. Differential Geom **17(2)** (1982), 255306.

32. R.S. Hamilton, *Non-singular solutions of the Ricci flow on three-manifolds*, Comm. Anal. Geom. **7(4)** (1999), 695-729.

33. M.Hirsch, *Differential Topology, Graduate Text in Mathematics 33*, Springer-Verlag, New-York, Heidelberg, Berlin, 1976.

34. D.Hoffman and J.Spruck, *Sobolev and isoperimetric inequalities for riemannian submanifolds*, Communications on Pure and Applied Mathematics **27, no 6** (1974), 715727.

35. P.Kronheimer and T.Mrowka, *Monopoles and Three-Manifolds* (2007), Cambridge University Press, New Mathematical Monographs.

36. W.H.Meeks and S.T.Yau, *The existence of embedded minimal surfaces and the problem of uniqueness*, Math.Z. **179** (1982), 151-168.

37. W.H.Meeks and S.T.Yau, *The classical Plateau problem and the topology of three dimensional manifolds*, Topology **21** (1982), 409-442.

38. J.Morgan and G.Tian, *Ricci Flow and the Poincare Conjecture*, vol. 3, Clay Mathematics Monograph, AMS and Clay Institute, 2007.

39. C.B.Morrey, *Multiple integrals in the calculus of variations*, New York, Springer-Verlag, 1966.

40. J.Nelson, *Applications of automatic transversality in contact homology, A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Mathematics) at the University of Wisconsin* (2013).

41. L.I.Nicolaescu, *Lectures on the Geometry of Manifolds*, World Scientific, 2007.

42. T.Parker and C.H.Taubes, *On Witten's proof of the Positive Energy Theorem*, Comm. Math.Phys **84 223-238** (1982).

43. G.Perelman, *Finite extinction time for the solutions to the Ricci flow on certain three-manifolds*, math.DG/0307245.

44. G.Perelman, *Ricci flow with surgery on three-manifolds*, math.DG/03031109.

45. Y.Rebai, *Study of the singular Yamabe problem in some bounded domain of R^n*, Adv. Nonlinear Stud **6, no 3** (2006), 437-460.

46. E.Sandier and S.Serfaty, *Vortex Patterns in Ginzburg-Landau Minimizers, XVIth International Congress Om Mathematical Physics. Held 3-8 August 2009 in Prague, Czech Republic. Edited by Pavel Exner (Doppler Institute, Prague, Czech Republic)* (2010), World Scientific Publishing, 246-264.

47. R.Schoen, *Conformal Deformation of a Riemannian metric to one of constant scalar curvature*, Jour.Diff.Geometry **20** (1984), 479-495.

48. R.Schoen, *The existence of weak solutions with prescribed singular behavior for a conformally invariant scalar equation.*, Comm.Pure Appl.Math **41, no3** (1988), 317392.

49. R.Schoen, *Estimates for stable minimal surfaces in three-dimensional manifolds*, Annals of Mathematical Studies, Princeton University Press **103** (1983), 111-126.

50. R.Schoen, *Variational Theory for the Total Scalar Curvature Functional for Riemannian Metrics and Related Topics* **1365** (1987), Topics in Calculus of Variations, Lecture Notes in Mathematics , Montecatini Terme, 120-154.

51. R.Schoen and L.Simon, *Regularity of stable Minimal surfaces*, Comm Pure and Applied Maths **XXXIV 741-797** (1981).

52. R.Schoen and L.Simon, *A New Proof of the Regularity Theorem for Rectifiable Currents which Minimize Parametric Elliptic Functionals*, Indiana University Mathematics Journal **31, no 3415-435** (1982).

53. R.Schoen and L.Simon, *Regularity and Singularity Estimates on Hypersurfaces minimizing Parametric Elliptic Variational Integrals I*, Acta Mathematica **139, no 3-4 217-265** (1977).

54. R.Schoen, L.Simon and S.T.Yau, *Curvature Estimates for Minimal Hyper-surfaces*, Acta Math. **134** (1975), 275-288.

55. R.Schoen and S.T.Yau, *On the proof of the positive mass conjecture in general relativity*, Comm.Math.Phys/vol 65, no 1 (1979), 45-76.

56. R.Schoen and S.T.Yau, *Conformally flat manifolds, Kleinian groups and scalar curvature*, Invent. Math **92 no 1** (1988), 4771.

57. L.Simon, *Lectures on Geometric Measure Theory*, vol. 3, Proc.Centre Math.Anal.Austral.Nat.Univ., 1983.

58. I.M.Singer, *Some remarks on the Gribov ambiguity*, Comm.Math.Phy **60** (1978), 7-12.

59. C. H. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture*, Geom. Topol. **11** (2007), 2117–2202.

60. E.Witten, *A new proof of the positive energy theorem*, Comm.Math. Phys **80** (1981), 381-402.