ELSEVIER

Journal of Geometry and Physics 21 (1997) 136–148

JOURNAL OF GEOMETRY AND PHYSICS

Instantons and hypercontact structures. Part II

Augustin Banyaga [1]

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

Received 22 August 1995; revised 9 February 1996

Abstract

We find sufficient conditions for the components of an $SU(2)$ connection to form a hypercontact structure. The hypercontact structure obtained depends canonically on the connection. The result implies that the components of any 1-(anti) instanton form a hypercontact structure on S^7.

Subj. Class.: Differential geometry
1991 MSC: 53C12, 53C15
Keywords: 1-(anti) instanton; Contact forms; Contact structures; Moduli space; Hyperkaehler structures; Hypercontact structures; 3-Sasakian structures; Hyper Tashiro construction

1. Introduction

This paper is a complement to [4]. We construct here a canonical hypercontact structure underlying any 1-(anti) instanton. We also show that Geiges–Thomas remark that the basic instanton yields a hypercontact structure can be obtained as a consequence of the "hyper" Tashiro construction.

The notion of hypercontact structure was recently coined by Geiges and Thomas [9]; it generalizes the notion of 3-Sasakian structures, which have been known for a couple of decades, see for instance [12]. For a recent extensive study of 3-Sasakian structures, we refer to [5,6]. These structures are the odd versions of hyperkaehler structures. We refer to [2] for a short but deep introduction to hyperkaehler manifolds which have been known for some decades as well, but only regained interest after the discovery of their connection with gauge theory and supersymmetry in mathematical physics [11]. Recently Geiges and Thomas pointed out a connection between hypercontact structures and gauge theory.

[1] Supported in part by NSF grant DMS 94-03196.

PII S0393-0440(96)00012-5

An instanton is an (anti) self-dual connection of an $SU(2)$ bundle over the 4-dimensional sphere S^4. Such an instanton is determined, up to a gauge transformation by an integer, the instanton number. We call "k-instantons" those which correspond to the integer k. Geiges and Thomas proved that the components of the basic 1-(anti) instanton form a hypercontact structure on S^7 [8].

The starting point of the research reported here was an attempt to find a direct proof of that fact. In Section 3, we give such a proof. It is a consequence of the "hyper" version of the Tashiro construction [3]. In [4], we extended, supplemented and formalized Geiges–Thomas arguments into a theorem asserting that the components of any 1-(anti) instanton are contact forms. Here, we prove that actually they form a hypercontact structure. The underlying contact structure depends canonically on the instanton. Finally, we propose here a definition of "hypercontact structures" which slightly relaxes Geiges–Thomas' and seems more appropriate to contact geometry.

2. Hypercontact structures

Recall that a hyperkaehler structure on a riemannian manifold (M, g) is a set of three complex structures J_1, J_2, J_3 such that

$$g \circ J_i = g, \quad i = 1, 2, 3, \tag{2.1}$$

$$J_1 J_2 = -J_2 J_1 = J_3, \qquad J_3 J_1 = -J_1 J_3 = J_2, \qquad J_2 J_3 = -J_3 J_2 = J_1, \tag{2.2}$$

$$\mathrm{d}\omega_i = 0, \tag{2.3}$$

where ω_i are defined by $\omega_i(X, Y) = g(X, J_i Y)$ for all vector fields X, Y.

A hyperkaehler manifold must be $4n$-dimensional since its tangent space is a quaternionic vector bundle. It comes equipped with three symplectic forms ω_i.

A hypercontact structure on a $(4n + 3)$-dimensional manifold is the analogue of a hyperkaehler manifold. Roughly speaking, it consists of a set of three contact forms with a hyperkaehler structure on each fiber of the intersection of the three contact distributions.

Recall that a contact form on a $(2n + 1)$-dimensional manifold is a 1-form α such that $\alpha \wedge (\mathrm{d}\alpha)^n$ is everywhere non-zero. A contact structure on a $(2n + 1)$-dimensional manifold M is a hyperplane field $F \subset TM$ which is locally the kernel of (a locally defined) contact form. If there is a globally defined contact form α such that $\operatorname{Ker} \alpha = F$, then the contact structure is called a contact structure in the "narrow sense" and one says it is defined by the contact form α; the distribution F is called the contact distribution. Two contact forms α and α' define the same contact structure if and only if there exists a nowhere zero function λ such that $\alpha' = \lambda\alpha$. Problems of Mechanics deal directly with "contact forms", while Contact Geometry deals more with "contact structures". Geiges and Thomas gave the definition of hypercontact structures in terms of "contact forms". We propose in this paper a definition in terms of "contact structures".

Definition 2.1 (Geiges–Thomas [9]). A set of three contact forms $\{\alpha_1, \alpha_2, \alpha_3\}$ on a $(4n+3)$-dimensional manifold M is a hypercontact structure if there exists a riemannian metric g, three 1–1 tensor fields ϕ_i, three 1-forms η_i and three vector fields ξ_i on M such that:

$$\eta_i(\xi_j) = \delta_{ij}, \tag{2.4}$$

$$\phi_i \xi_j = \epsilon_{ijk}\xi_k, \tag{2.5}$$

$$\eta_i \circ \phi_j = \epsilon_{ijk}\eta_k, \tag{2.6}$$

$$\phi_i\phi_j(X) = -\delta_{ij}X + \eta_j(X)\xi_i + \epsilon_{ijk}\phi_k X, \tag{2.7}$$

$$g(X, Y) = g(\phi_i X, \phi_i Y) + \eta_i(X)\eta_i(Y), \tag{2.8}$$

$$g(X, \phi_i Y) = \mathrm{d}\alpha_i(X, Y) \tag{2.9}$$

for all vector fields X, Y.

Here ϵ_{ijk} is zero when all the symbols are not distinct, and if they are, it is equal to the signature of the corresponding permutation of the integers 1, 2, 3. If furthermore, we have:

$$\alpha_i = \eta_i, \tag{2.10}$$

$$L_{\xi_i} g = 0, \tag{2.11}$$

$$[\xi_i, \xi_j] = 2\epsilon_{ijk}\xi_k, \tag{2.12}$$

we say that $\{\alpha_1, \alpha_2, \alpha_3\}$ is a 3-Sasakian structure.

Remark 2.2.

(1) In Eq. (2.11), L_X stands for the Lie derivative and Eq. (2.11) says that the vector fields are Killing with respect to the riemannian metric g. This is a very strong condition which makes 3-Sasakian structures rigid. Hypercontact structures are more flexible: for instance they admit connected sums [9].

(2) Conditions (2.7)–(2.9) say that on the intersection of the contact distribution, (ϕ_i, g) form hyperkaehler structures in each fiber.

(3) The definition given above is in terms of "contact forms". We propose here to recapture the essential notions by using "contact structures". Perhaps it is necessary to assume that the contact structures are defined by global contact forms.

Definition 2.3. A set of three contact structures F_1, F_2, F_3 (defined by global contact forms) on a $(4n+3)$-dimensional manifold M is said to be a hypercontact structure if:

(i) $H = F_1 \cap F_2 \cap F_3$ is a $4n$-dimensional distribution.

(ii) There exist a riemannian metric g and three endomorphisms ϕ_i of H satisfying the quaternionic identities (2.2) and such that $g(\phi_i X, \phi_i Y) = g(X, Y)$ for all sections X, Y of H,

(iii) For any choice of contact forms α_i representing F_i, there are positive functions λ_i such that $\mathrm{d}\alpha_i(X, Y) = \lambda_i g(X, \phi_i Y)$ for all sections X, Y of H.

(1) This definition is independent of the choice of the contact forms α_i since on the contact distribution, $\mathrm{d}\alpha_i$ defines a conformal symplectic structure depending only on the contact structure H_i.

(2) There exist contact forms α_i, defining H_i such that $\mathrm{d}\alpha_i(X, Y) = g(X, \phi_i Y)$ for all sections X, Y of H. Indeed, if α_i' are contact forms like in (iii), such that $\mathrm{d}\alpha_i'(X, Y) = \lambda_i g(X, \phi_i Y)$ for all sections X, Y of H, then setting $\alpha_i = (1/\lambda_i)\alpha_i'$, we get $\mathrm{d}\alpha_i(X, Y) = g(X, \phi_i Y)$ for all sections X, Y of H.
(3) The Geiges–Thomas definition needs more assumptions. Going from our definition to theirs will involve non-canonic choices, like an orthonormal basis of the orthogonal complement to H, and a set of three contact forms α_i representing the contact structures H_i and such that $\mathrm{d}\alpha_i(X, Y) = g(X, \phi_i Y)$ for all sections X, Y of H. Assume now the orthonormal complement V of H admits three orthonormal sections ξ_1, ξ_2, ξ_3 and let η_i be the dual forms of ξ_i. Now extend the definition of the endomorphisms ϕ_i of H by Eq. (2.5) in the Geiges–Thomas definition. It is easy to verify that with the above choices of $\alpha_i, \eta_i, \phi_i, g$, Eqs. (2.4)–(2.9) hold, so we get a hypercontact structure in the Geiges–Thomas sense. In practice, the ξ_i will be proportional to the Reeb fields of α_i.

This procedure is more geometric and constructive. We will use it in the proof of Theorem 2.

2.1. *Hypersurfaces of hypercontact type in hyperkaehler manifolds*

A classical way of getting contact manifolds is to obtain them as hypersurfaces of contact type in symplectic manifolds [16]. Let $\rho : M \to P$ be the inclusion of a hypersurface M into a symplectic manifold (P, Ω). If there exists a vector field V on a neigbourhood of M, which is transverse to M, and which is a Liouville vector field, meaning that $L_V \Omega = \lambda \Omega$, for some positive function λ, then $\alpha = \rho^*(i(V)\Omega)$ is a contact form on M such that $\rho^*\Omega = (1/k)\,\mathrm{d}\alpha$, where $k = \lambda \circ \rho$.

Here L_V stands for the Lie derivative in the direction of the vector field V, and $i(V)$ is the interior product with V.

We can try to do the same with a hypersurface in a hyperkaehler manifold. The point is that there must be a transverse vector field which is Liouville with respect to the three symplectic forms [9].

We consider here a special case which is particularly nice: the "hyper" Tashiro construction.

First let us recall the Tashiro construction. (See for instance [3].)

Let (P, G, J) be an almost hermitian manifold: J is an almost complex structure and G is a riemannian metric such that $G \circ J = G$. Let $\rho : M \to P$ be an oriented hypersurface. The unit "outward" normal vector to M is well defined: it is a section C of the normal bundle such that $G(x)(C_x, C_x) = 1$ and $G(x)(C_x, X_x) = 0$ for all x in M, and all tangent vector X_x to M at x. Since $G_x(J_x C_x, C_x) = G_x(J_x^2 C_x, J_x C_x) = -G_x(C_x, J_x C_x)$, we see that $G_x(J_x C_x, C_x) = 0$, hence $J_x C_x$ is tangent to M. We thus have defined a vector field ξ on M such that

$$(\rho_* \xi)_x = -J_x C_x, \tag{2.13}$$

which is called the characteristic vector field of the hypersurface M [3]. For any vector field X on M, we let $\eta(X)$ be the norm of the projection of $J\rho_* X$ onto the C-direction, i.e.

$$\eta(X) = G(C, J\rho_* X). \tag{2.14}$$

Therefore $J\rho_* X - \eta(X)C$ is tangent to M. We have thus defined a vector field ϕX such that

$$\rho_*(\phi X) = J\rho_* X - \eta(X)C. \tag{2.15}$$

The 1-form η, the vector field ξ, the 1–1 tensor field ϕ, and the restriction $g = \rho^* G$ of the metric G to M satisfy the following identities:

$$\eta(\xi) = 1, \tag{2.16}$$

$$\phi\xi = 0, \tag{2.17}$$

$$\phi^2 X = -X + \eta(X)\xi, \tag{2.18}$$

$$g(X, Y) = g(\phi X, \phi Y) + \eta(X)\eta(Y). \tag{2.19}$$

As a consequence of (2.15), (2.16) and (2.19), we have

$$\eta \circ \phi = 0; \qquad g(X, \xi) = \eta(X). \tag{2.20}$$

The two skew symmetric forms: $\Omega(X, Y) = G(X, JY)$ on P and $\Phi(U, V) = g(U, \phi V)$ on M have maximum rank. Hence the kernel of Φ is 1-dimensional and is spanned by ξ. Moreover, $\rho^*\Omega = \Phi$. We have:

$$\eta \wedge \Phi^n \neq 0, \tag{2.21}$$

$$\rho^*[i(C)\Omega] = \eta. \tag{2.22}$$

If now we assume that Ω is closed (i.e. a symplectic form) and that C is Liouville, then η is a contact form.

Now suppose we have a hyperkaehler manifold $(P^{4n+4}, G, J_1, J_2, J_3)$ and let ρ: $M^{4n+3} \to P^{4n+4}$ be a hypersurface such that the unit normal vector field along M is Liouville for the three symplectic forms $\Omega_i(X, Y) = G(X, J_i Y)$, $i = 1, 2, 3$, i.e. $L_C \Omega_i = \lambda_i \Omega_i$. Then the 1-forms η_i are contact forms and we have

$$\mathrm{d}\eta_i = \mu_i \rho^* \Omega = \mu_i \Phi, \tag{2.23}$$

where $\mu_i = (\lambda_i \circ \rho)$.

Lemma 2.4. *Suppose that $\mu_1 = \mu_2 = \mu_3 = \mu$ is a constant, then the contact forms $(\alpha_1, \alpha_2, \alpha_3)$, where $\alpha_i = (1/\mu)\eta_i$, are a hypercontact structure.*

Proof. We have the three vector fields ξ_k defined by $\rho_*\xi_k = -J_k C$, the three 1-forms η_k defined by $\eta_k(X) = G(C, J_k\rho_* X)$ and three 1–1 tensor fields ϕ_k defined by $\rho_*\phi_k X =$

$J_k\rho_*X - \eta_k(X)C$. We are now going to prove that the triplet (η_k, ϕ_k, ξ_k) satisfy conditions (2.4)–(2.9) of the definition of a hypercontact structure.

The hypothesis of the lemma guarantees that condition (2.9) holds. Condition (2.8) was verified by construction: it is the identity (2.19) in the Tashiro construction. To verify (2.4), we write

$$\eta_i(\xi_j) = G(C, J_i(\rho_*\xi_j) = -G(C, J_iJ_jC) = \delta_{ij}.$$

To check (2.5), we write

$$\begin{aligned}\eta_i(\phi_jX) &= G(C, J_i\rho_*\phi_jX) = G(C, J_i(J_j\rho_*X - \eta_j(X)C))\\ &= G(C, J_iJ_j\rho_*X) - \eta_j(X)G(C, J_iC)\\ &= G(C, \epsilon_{ijk}J_k\rho_*X) = \epsilon_{ijk}\eta_k(X).\end{aligned}$$

To verify (2.6), just observe that $\rho_*(\phi_i\xi_j) = J_i(\rho_*\xi_j) - \eta_i(\xi_j)C = -J_iJ_jC - \delta_{ij}C$. The only non-trivial relation is (2.7). For $i=j$ this is the relation (2.18) in the Tashiro construction. We verify that

$$\phi_1\phi_2X = \eta_2(X)\xi_1 + \phi_3X,$$

and let the reader establish the other relations by circular permutations.

$$\begin{aligned}\rho_*(\phi_1\phi_2X) &= J_1(\rho_*\phi_2X) - \eta_1(\phi_2X)C\\ &= J_1(J_2\rho_*X - \eta_2(X)C) - \eta_3(X)C\\ &= J_1J_2\rho_*X + \eta_2(X)\rho_*\xi_1 - \eta_3(X)C\\ &= (J_3\rho_*X - \eta_3(X)C) + \eta_2(X)\rho_*\xi_1\\ &= \rho_*(\phi_3X + \eta_2(X)\xi_1).\end{aligned}$$

The proof of the lemma is now complete. □

Example 2.5. Let $P = \mathbb{R}^{4n} = \mathbb{R}^4 \times \mathbb{R}^4 \times \cdots \times \mathbb{R}^4$ and $\mathcal{J}_k = (J_k, \ldots, J_k)$, where J_k are the following complex structures on $\mathbb{R}^4$:

$$\begin{aligned}&J_1\partial_1 = \partial_2, \qquad J_1\partial_3 = \partial_4, \qquad J_2\partial_1 = \partial_3,\\ &J_2\partial_2 = \partial_4, \qquad J_3\partial_1 = -\partial_4, \qquad J_3\partial_2 = \partial_3.\end{aligned}$$

Let G be the standard dot product on $\mathbb{R}^{4n}$. Then $(P, G, \mathcal{J}_i, i = 1, 2, 3)$ is a hyperkaehler manifold. The sphere S^{4n-1} is the set $f^{-1}(1)$ where $f(x) = |x|^2$. Let C be half the gradient of f. Then C is a Liouville field with $L_C\Omega_k = 2\Omega_k$. By the lemma, the 1-forms η_k and $\alpha_k = \frac{1}{2}\eta_k$ are contact forms and $(\alpha_1, \alpha_2, \alpha_3)$ are a hypercontact structure.

Writing points in $\mathbb{R}^{4n}$ as $(x_1^i, x_2^i, x_3^i, x_4^i)_{i=1,\dots,n}$, the contact forms α_k read:

$$\alpha_1 = \sum_{i=1}^{n} x_2^i \,\mathrm{d}x_1^i - x_1^i \,\mathrm{d}x_2^i + x_4^i \,\mathrm{d}x_3^i - x_3^i \,\mathrm{d}x_4^i,$$

$$\alpha_2 = \sum_{i=1}^{n} x_3^i \,\mathrm{d}x_1^i - x_1^i \,\mathrm{d}x_3^i + x_2^i \,\mathrm{d}x_4^i - x_4^i \,\mathrm{d}x_2^i,$$

$$\alpha_3 = \sum_{i=1}^{n} x_4^i \,\mathrm{d}x_1^i - x_1^i \,\mathrm{d}x_4^i + x_3^i \,\mathrm{d}x_2^i - x_2^i \,\mathrm{d}x_3^i.$$

3. Instantons

An $SU(2)$ bundle $\pi : M \to S^4$ over S^4 is determined by an integer k_π, which is the degree of the transition map of two sections defined over U, V where $U = S^4 - \{p\}$ and $V = S^4 - \{q\}$. The transition is a map from $U \cap V \approx S^3$ into $SU(2) = S^3$. Self-dual and anti-self-dual connections on an $SU(2)$ bundle over S^4 with characteristic integer k_π are called k_π-instantons and anti-instantons. These are the minima of the Yang–Mills functional. See Atiyah's collected work [1] for a comprehensive foundation of gauge theory. Atiyah–Drinfeld–Hitchin–Manin (ADHM) gave a complete construction of instantons [1]. For $k_\pi = 1$ they proved that any 1-instanton is gauge equivalent to an instanton α such that $\mu^*\alpha = A(x)$ is given by the following expression:

$$\mu^*\alpha = \mathrm{Im}\left(\frac{(x-a)\,\mathrm{d}\bar{x}}{\lambda^2 + |x-a|^2}\right),$$

where $a \in \mathbb{H}$ is a quaternionic parameter and λ is a positive real number, and $\mu : \mathbb{R}^4 \to S^7$ is the following section over $S^4 - (0,0,0,0,1) \approx \mathbb{R}^4$:

$$\mu(x) = \frac{(x, 1)}{(1 + |x|^2)^{1/2}}.$$

The instanton corresponding to the case $a = 0$, $\lambda = 1$ is called the basic instanton. Its potential is

$$A(x) = \mathrm{Im}\left(\frac{x\,\mathrm{d}\bar{x}}{1 + |x|^2}\right).$$

Anti-instantons have exactly the same description, modulo putting the bars over x's instead of $\mathrm{d}x$'s. Geometrically, this corresponds to a change of the orientation of the bundle.

First a review of some quaternionic notations. The field $\mathbb{H}$ of quaternions is the set $\{x = x_1 + x_2\mathrm{i} + x_3\mathrm{j} + x_4\mathrm{k}, x_i \in \mathbb{R}\}$ where $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$ and $\mathrm{ij} = -\mathrm{ji} = \mathrm{k}$; $\mathrm{jk} = -\mathrm{kj} = \mathrm{i}$; $\mathrm{ki} = -\mathrm{ik} = \mathrm{j}$. We naturally identify $\mathbb{H}$ with $\mathbb{R}^4$ and with $\mathbb{C}^2$. Writing $x = x_1 + x_2\mathrm{i} + x_3\mathrm{j} + x_4\mathrm{k} = z_1 + z_2\mathrm{j}$ where $z_1 = x_1 + x_2\mathrm{i}$, $z_2 = x_3 + x_4\mathrm{i}$ establishes an identification of $\mathbb{H}$ and $\mathbb{C}^2$. The conjugate $\bar{x}$ of a quaternion x is $x_1 - x_2\mathrm{i} - x_3\mathrm{j} - x_4\mathrm{k}$ and $x\bar{x} = \bar{x}x = |x|^2$. Also $\mathbb{H}$ can be viewed as the set of 2×2 complex matrices: $x = z_1 + z_2\mathrm{j}$ corresponds to the matrix

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix},$$

the determinant of which is the square norm of x. Therefore $SU(2)$ is the group of norm-1 quaternions, i.e. a sphere S^3. Its Lie algebra $su(2)$ is the set of skew hermitian matrices with zero trace. The Pauli matrices

$$\tau_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad \tau_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \tau_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

form a basis of $su(2)$. Their commutation relations are

$$[\tau_1, \tau_2] = 2\tau_3, \qquad [\tau_1, \tau_3] = -2\tau_2, \qquad [\tau_2, \tau_3] = 2\tau_1.$$

Hence $su(2)$ is isomorphic with the imaginary part $\{x_2 i + x_3 j + x_4 k\}$ of $\mathbb{H}$: we identify τ_1 with i, τ_2 with j and τ_3 with k.

Now $S^7 = \{(p, q) \in \mathbb{H}^2, |p|^2 + |q|^2 = 1\}$, and S^4 is the $\mathbb{H}$ projective line, i.e. the set of equivalence classes $[p, q]$ of elements in $\mathbb{H}^2 - \{0\}$: $(p, q) \sim (p', q')$ iff $p = rp'$, $q = rq'$ for some $r \in \mathbb{H} - \{0\}$.

The bundle map π of the tautological bundle $\pi : S^7 \to S^4$ assigns to $(p, q) \in S^7$ the equivalence class $[p, q] \in S^4$. This is a principal $SU(2)$ bundle with Pontryagin number $k = +1$. It is easy to see that

$$\alpha(p, q) = \operatorname{Im}(p \, \mathrm{d}p + q \, \mathrm{d}\overline{q})$$

is a connection such that

$$\mu^* \alpha = \operatorname{Im}\left(\frac{x \, \mathrm{d}\overline{x}}{1 + |x|^2} \right),$$

where $\mu : \mathbb{R}^4 \to S^7$ is the section over $S^4 - (0, 0, 0, 0, 1) \approx \mathbb{R}^4$:

$$\mu(x) = \frac{(x, 1)}{(1 + |x|^2)^{1/2}}.$$

In other words, α is the basic instanton [1]. See [13, pp. 100–104].

Setting $p = x_1 + x_2 i + x_3 j + x_4 k$, $q = y_1 + y_2 i + y_3 j + y_4 k$, and $\alpha = (\alpha_1) i + (\alpha_2) j + (\alpha_3) k$, we have:

$$\begin{aligned}
\alpha_1 &= x_2 \, \mathrm{d}x_1 - x_1 \, \mathrm{d}x_2 + x_4 \, \mathrm{d}x_3 - x_3 \, \mathrm{d}x_4 + y_2 \, \mathrm{d}y_1 - y_1 \, \mathrm{d}y_2 + y_4 \, \mathrm{d}y_3 - y_3 \, \mathrm{d}y_4, \\
\alpha_2 &= x_3 \, \mathrm{d}x_1 - x_1 \, \mathrm{d}x_3 + x_2 \, \mathrm{d}x_4 - x_4 \, \mathrm{d}x_2 + y_3 \, \mathrm{d}y_1 - y_1 \, \mathrm{d}y_3 + y_2 \, \mathrm{d}y_4 - y_4 \, \mathrm{d}y_2, \\
\alpha_3 &= x_4 \, \mathrm{d}x_1 - x_1 \, \mathrm{d}x_4 + x_3 \, \mathrm{d}x_2 - x_2 \, \mathrm{d}x_3 + y_4 \, \mathrm{d}y_1 - y_1 \, \mathrm{d}y_4 + y_3 \, \mathrm{d}y_2 - y_2 \, \mathrm{d}y_3.
\end{aligned}$$

These forms are exactly those we constructed before by the “hyper” Tashiro construction, and we showed they form a hypercontact structure.

Therefore, we have given a direct proof of the following result (see also [4]):

Theorem 3.1 (Geiges–Thomas [8]). *The components of the basic instanton form a hypercontact structure on S^7.*

Let us now give a brief invariant description of instantons following [7] or [14]. Let η_{S^7} be the $\mathbb{H}$-bundle associated with the tautological principal bundle π. Let D be the connection induced on η_{S^4} by the connection α. The double cover $SL(2, \mathbb{H})$ of the group $SO(5, 1)$ of conformal transformations of S^4 acts on η_{S^4}. Since the Yang–Mills equations are conformally invariant, the induced connection g^*D, for g in $SL(2, \mathbb{H})$, is self-dual. A difficult theorem of Atiyah et al. [1] asserts that we get this way all the gauge equivalences of instantons. The formulas for instantons we wrote before are nothing else than the local description of this fact. Since $SU(2)$ acts freely on η_{S^7}, each connection g^*D comes from a unique connection α_g on the principal bundle [15]. These are the instantons we consider on S^7.

The main result of this paper is the following generalization of the Geiges–Thomas theorem.

Theorem 3.2. *The components of any* 1*-(anti) instanton form a hypercontact structure on* S^7.

Remark 3.3. We insist that this theorem is not trivial. Indeed, although we proved that the components of α form a hypercontact structure, this does not imply that the components of α_g form a hypercontact structure as well. Observe that the group $SL(2, \mathbb{H})$ does not act on S^7, hence α_g is not the pull-back of some connection on π by a diffeomorphism of S^7.

Theorem 3.2 is a consequence of the following generalization of a result of [4].

Theorem 3.4. *Let* $\pi: P \to M$ *be a principal* $SU(2)$ *bundle and* α *a connection with curvature* Ω *and let* α_i, Ω_i, $i = 1, 2, 3$, *be the components of* α *and* Ω *along the Pauli matrices (basis of* $su(2)$*). Suppose there is a family of sections* $\sigma_j : U_j \to P$ *trivializing the bundle (here* $\{U_j\}$ *is an open cover over which the bundle is trivial), and smooth nowhere vanishing functions* ν_j *on* U_j *such that* $\{\nu_j \sigma_j^* \Omega_i\}$, $i = 1, 2, 3$, *form a hyperkaehler structure on* U_j, *then* $\{\alpha_1, \alpha_2, \alpha_3\}$ *form a hypercontact structure on* P.

In [4, Theorem 2] we only proved that the three forms $\{\alpha_i\}$ are contact forms. Here we show how to put together an underlying hypercontact structure. This hypercontact structure happens to be defined canonically by the connection.

Proof of Theorem 3.4. We need to recall some notations and part of the proof of Theorem 2 of [4].

Let $\{U\}$ be a trivializing open cover like in the hypothesis of the theorem, and let $\sigma : U \to P$ be a section and ν a smooth nowhere zero function on U such that $\{w_i = \nu\sigma^*\Omega_i\}$, $i = 1, 2, 3$, form a hypersymplectic structure, i.e. there exists a riemannian metric g on U, three almost complex structures J_i on U satisfying the quaternionic identities (see definition above) and such that $g(J_i X, Y) = w_i(X, Y)$ and $g \circ J_i = g$.

We denote by $H \subset T(P)$ the horizontal space, i.e. the kernel of α and by $G = \pi^*g$ the pull-back of the metric g on $P_U = \pi^{-1}(U)$. If X is a vector field on P, we denote by X_{h} its horizontal component. If X is horizontal, then $(\sigma_*(\pi_* X))_{\mathrm{h}} = X$ and $\Omega(X_{\mathrm{h}}, \cdot) = \Omega(X, \cdot)$ since Ω vanishes on vertical vectors.

Let now X, Y be two horizontal vector fields on P at $\sigma(x), x \in U$:

$$\begin{aligned}\Omega_i(\sigma(x))(X, Y) &= \Omega_i(\sigma(x))((\sigma_*\pi_*X)_{\rm h}, (\sigma_*\pi_*Y)_{\rm h}) \\ &= \Omega_i(\sigma(x))(\sigma_*\pi_*X, \sigma_*\pi_*Y) \\ &= (\sigma^*\Omega_i)(x)(\pi_*X, \pi_*Y) \\ &= (1/\nu)g(J_i\pi_*X, \pi_*Y).\end{aligned}$$

This shows that Ω_i are non-degenerate at $H_{\sigma(x)}, x \in U$, since π_* is an isomorphism between the horizontal space at $\sigma(x)$ and the tangent space at $x \in U$.

Any other point $p \in P_U$ has the form $p = \sigma(x) \cdot a = R_a(\sigma(x))$ for some $a \in SU(2)$. If X_p is a horizontal vector field at $p = \sigma(x) \cdot a$, i.e. $X_p \in H_p$, then $X_p = (R_a)_*X_{\sigma(x)}$. Hence for $X_p, Y_p \in H_p$, we have

$$\begin{aligned}\Omega(p)(X_p, Y_p) &= \Omega(R_a(\sigma(x)))((R_a)_*X_{\sigma(X)}, (R_a)_*Y_{\sigma(x)}) \\ &= (R_a^*\Omega)(\sigma(x))(X_{\sigma(x)}, Y_{\sigma(x)}).\end{aligned}$$

But the curvature form satisfies $R_a^*\Omega = ad_{a^{-1}}(\Omega) = a\Omega a^{-1}$. Let (μ_{ij}) be the matrix of $ad_{a^{-1}} : su(2) \to su(2)$ within the basis τ_1, τ_2, τ_3, then:

$$\begin{aligned}\Omega_i(p)(X_p, Y_p) &= \sum_{j=1}^{3} \mu_{ij}\Omega_j(\sigma(x))(X_{\sigma(x)}, Y_{\sigma(x)}) \\ &= \sum_{j=1}^{3} \mu_{ij}(1/\nu)g(J_j\pi_*X, \pi_*Y) = (1/\nu)g(\Phi_i\pi_*X, \pi_*Y),\end{aligned}$$

where $\Phi_i = \sum_{j=1}^{3} \mu_{ij}J_j$. Since $ad_{a^{-1}}$ preserves the natural inner product: $(m, n) = -\frac{1}{2}\mathrm{tr}(m.n)$, the matrix (μ_{ij}) is an orthogonal matrix. This implies that the 1–1 tensors defined on U satisfy the quaternionic identities since the J_i's did. In particular they define complex structures on U depending on $a \in SU(2)$. The equation

$$\Omega_i(p)(X_p, Y_p) = (1/\nu)g(\Phi_i\pi_*X, \pi_*Y) \tag{3.1}$$

shows that Ω_i are non-degenerate at the horizontal distribution at $\sigma(x) \cdot a$.

Furthermore, since the metric g over U was compatible with the complex structures J_i, i.e. $g \circ J_i = g$, and since (μ_{ij}) is an orthogonal matrix, we must have

$$g(\Phi_iX, \Phi_iY) = g(X, Y) \tag{3.2}$$

for all horizontal vector fields X, Y.

Let us now check that the mappings $\mathrm{d}\alpha_i^b : TP \to T^*P$ defined by: $\mathrm{d}\alpha_i^b(X)(Y) = \mathrm{d}\alpha_i(X, Y)$ map bijectively H to H^*. It is sufficient to check that $\mathrm{d}\alpha_i(X, \xi_k) = 0$ for all ξ_k and all sections X of H. The structural equation says that $\Omega_i = \mathrm{d}\alpha_i + \alpha_j \wedge \alpha_l$.
Hence

$$\mathrm{d}\alpha_i(X, \xi_k) = (\Omega_i - \alpha_j \wedge \alpha_l)(X, \xi_k) = 0,$$

since $i(\xi_k)\Omega_i = 0$, ξ_k being a vertical vector field and $\alpha_l(X) = 0$ since X is horizontal. We can now define three endomorphisms of H as follows:

$$\phi_1 = -(\mathrm{d}\alpha_2^b)^{-1} \circ (\mathrm{d}\alpha_3^b), \tag{3.3}$$

$$\phi_2 = -(\mathrm{d}\alpha_3^b)^{-1} \circ (\mathrm{d}\alpha_1^b), \tag{3.4}$$

$$\phi_3 = -(\mathrm{d}\alpha_1^b)^{-1} \circ (\mathrm{d}\alpha_2^b). \tag{3.5}$$

These three endomorphisms satisfy the quaternionic identities. Indeed, we saw that over each open set U, the forms Ω_i were given by (3.1). Hence over that open set, $\Phi_i = (g_v^b)^{-1} \circ (\mathrm{d}\alpha_i)^b$, where g_v stands for the metric $(1/v)g$ in Eq. (3.1) and $g_v^b : TP \to T^*P$ is $g_V^b(X)(Y) = g_v(X, Y)$. Consequently, we see that the only locally defined $\Phi_i^{-1} \circ \Phi_j = (((\mathrm{d}\alpha_i)^b)^{-1} \circ g_v^b)(((g_v^b)^{-1}) \circ \mathrm{d}\alpha_j^b)$ coincides with the globally defined $(\mathrm{d}\alpha_i^b)^{-1} \circ \mathrm{d}\alpha_j^b$. Since the locally defined object satisfy the quaternionic identities on H so do the endomorphisms ϕ_i. Now we define a metric on H by setting

$$g^b = (\mathrm{d}\alpha_1^b) \circ \phi_1^{-1} = (\mathrm{d}\alpha_2^b) \circ \phi_2^{-1} = (\mathrm{d}\alpha_3^b) \circ \phi_3^{-1}. \tag{3.6}$$

This is positive definite since it is so over each open set where it is equal to g_v. We now extend ϕ_i and g in the vertical direction by

$$\phi_i(\xi_j) = \epsilon_{ijk}\xi_k, \qquad g(\xi_i, \xi_j) = \delta_{ij}, \qquad g(X, \xi_k) = 0 \tag{3.7}$$

for any horizontal vector field X. Finally, we get the three 1-forms η_k by setting

$$\eta_k(X) = g(X, \xi_k). \tag{3.8}$$

Let us now verify that the set $(\alpha_k, \eta_k, \xi_k, \phi_k, g)$ verify the conditions (2.4)–(2.9) (with the same notations), so the set $(\alpha_1, \alpha_2, \alpha_3)$ form a hypercontact structure.

Conditions (2.4) and (2.5) come straight from (3.7) and (3.8). Condition (2.9) results from (3.2) and the definition of g in (3.6). Condition (2.6) and (2.7) are immediate if we compute both sides of the equations on vertical and horizontal vector fields. Condition (2.8) is less trivial. Let us verify it. Any vector field X decomposes as $X = X_{\mathrm{H}} + b_1\xi_1 + b_2\xi_2 + b_3\xi_3$, where X_{H} is horizontal and $a_i = g(X, \xi_i) = \eta_i(X)$. Now if $Y = Y_{\mathrm{H}} + b_1 + b_2 + b_3$, then $g(X, Y) = g(X_{\mathrm{H}}, Y_{\mathrm{H}}) + a_1b_1 + a_2b_2 + a_3b_3$. Since $\phi_1 X = \phi_1 X_{\mathrm{H}} + a_2\xi_3 - a_3\xi_2$, we have $g(\phi_1 X, \phi_1 Y) = g(\phi_1 X_{\mathrm{H}}, \phi_1 Y_{\mathrm{H}}) + a_2b_2 + a_3b_3$. By (2) $g(\phi_1 X_{\mathrm{H}}, \phi_1 Y_{\mathrm{H}}) = g(X_{\mathrm{H}}, Y_{\mathrm{H}})$. Therefore,

$$\begin{aligned} g(\phi_1 X, \phi_1 Y) &= g(X_{\mathrm{H}}, Y_{\mathrm{H}}) + a_2b_2 + a_3b_3 = g(X, Y) - a_1b_1 \\ &= g(X, Y) - \eta_1(X)\eta_1(Y). \end{aligned} \tag{3.9}$$

This proves (2.8) for $i = 1$, but exactly the same argument works for $i = 2, 3$. The proof of Theorem 3.4 is now complete. □

4. Remarks

(1) Eqs. (3.3)–(3.7) show that the hypercontact structure $(\sigma_k, \eta_k, \xi_k, \phi_k, g)$ is canonically determined by the connection α. Recall that the vector fields ξ_k are the fundamental vector fields defined by the Pauli matrices (basis of $su(2)$, see [4]).

(2) In Theorem 3.4 we have dealt with the problem of showing that three contact forms are a hypercontact structure in a very special case. The general problem to show that three contact forms are a hypercontact structure seems very elusive. Geiges and Thomas have proved the existence of three linearly independent contact forms on any compact 2-connected 7-manifold [9]. It is an open question whether these form a hypercontact structure. In the next section, we give a necessary condition for three contact forms to form a hypercontact structure: they must represent equivalent contact structures.

(3) In the definition of a hypercontact structure the ingredients are tied up with strong relations. This suggests that they are not independent. For instance it is well known that in case of hyperkaehler manifolds (the even dimensional version of hypercontact structures), the riemannian metric is determined by the kaehler forms, which also determine the three complex structures. We just proved that in the hypercontact case, the riemannian metric is determined by the three contact forms, and so are the restrictions of the endomorphisms ϕ_i to the horizontal distribution.

(4) Also any linear combination of the three kaehler forms is again a kaehler form. If they had the same cohomology class, they would be all equivalent by Moser's theorem [14]. In [4], we proved the following result.

Theorem 4.1. *Let* $\{(\alpha_1, \alpha_2, \alpha_3), (\phi_i, \xi_i, \eta_i)_{i=1,2,3}\}$ *be a hypercontact structure on a riemannian manifold* (M, g) *such that* $\alpha_i(\xi_j) = 0, \forall i \neq j$. *Then the three contact forms* α_i *represent equivalent contact structures.*

Corollary 4.2. *Suppose the Reeb fields of three contact forms on a 3-dimensional manifold parallelize it. Then the three contact forms represent equivalent contact structures. In particular the Reeb fields of three contact forms on a 3-manifold, one of them being tight and another being overtwist, are not everywhere linearly independent.*

Proof. Indeed, on a 3-manifold, three contact forms, with Reeb field everywhere linearly independent form a hypercontact structure [9]. □

Corollary 4.3. *A necessary condition for three contact forms* α_i *with Reeb fields* ξ_i *such that* $\alpha_i(\xi_j) = 0, \forall_i \neq j$, *to be a hypercontact structure is that these contact forms determine equivalent contact structures.*

Acknowledgements

I would like to thank the University of Cambridge and the University of Montpellier for their hospitality while I was preparing this paper. Special thanks to C.B. Thomas for his interest in this work.

References

[1] M.F. Atiyah, Geometry of Yang–Mills fields, in: *Collected Work*, ed., M. Atiyah, Vol. 5 (Clarendon Press, Oxford, 1988) pp. 77–173.
[2] M.F. Atiyah, Hyperkaehler manifolds, in: *Complex Geometry and Analysis*, Lecture Notes in Mathematics, Vol. 1422 (Springer, Berlin, 1980) pp. 1–13.
[3] A. Banyaga, On characteristics of hypersurfaces in symplectic manifolds, *Proc. Symp. in Pure Mathematics* 54 (1993) 9–17.
[4] A. Banyaga, Instantons and hypercontact structures, J. Geom. Phys., to appear.
[5] C. Boyer, K. Galicki and B.M. Mann, Quaternionic reduction and Einstein manifolds, Comm. Anal. Geom. (1993) 229–279.
[6] C. Boyer, K. Galicki and B.M. Man, The geometry and topology of 3-Sasakian manifolds, J. Reine Angew. Math. (1994) 183–220.
[7] D.S. Freed and K.K. Uhlenbeck, *Instantons and Four-Manifolds* (MSRI Publication, Springer, Berlin, 1984).
[8] H. Geiges and C.B. Thomas, Sasakian 3-structure on the 7-sphere and the basic instanton, preprint.
[9] H. Geiges and C.B. Thomas, Hypercontact manifolds, J. London Math. Soc., to appear.
[10] J. Gray, Some properties of contact structures, Ann. Math. 69 (1959) 421–450.
[11] N.J. Hitchin, A. Karhede, U. Lindstrom and M. Rocek, Hyperkaehler metrics and supersymmetry, Comm. Math. Phys. 108 (1987) 535–589.
[12] M. Konishi, On manifolds with Sasakian 3-structure over quaternion Kaehler manifolds, Kodai Math. Sem. Rep. 26 (1975) 194–200.
[13] B.H. Lawson, The theory of gauge fields in four dimensions, *CBM Regional Conf. Series in Math.*, Vol. 58 (American Mathematical Society, Providence, RI, 1985).
[14] J. Moser, On the volume-element on manifolds, Trans. Amer. Math. Soc. 120 (1965) 280–296.
[15] S. Sternberg, *Lectures On Differential Geometry* (Prentice-Hall, Englewood Cliffs, NJ, 1964).
[16] A. Weinstein, On the hypothesis of Rabinowicz' periodic orbit theorem, J. Differential Equations 33 (1978) 353–358.