Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 345 (2007) 319–324
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Optimal Control

Optimality results in orbit transfer

Bernard Bonnard [a], Jean-Baptiste Caillau [b]

[a] *Institut de mathématiques de Bourgogne (UMR CNRS 5584), 9, avenue Savary, 21078 Dijon, France*
[b] *ENSEEIHT-IRIT (UMR CNRS 5505), 2, rue Camichel, 31071 Toulouse, France*

Received 14 November 2006; accepted after revision 27 July 2007

Available online 28 August 2007

Presented by Alain Bensoussan

Abstract

The objective of this Note is to present optimality results in orbital transfer. Averaging of the energy minimization problem is considered, and properties of the associated Riemannian metric are discussed. ***To cite this article: B. Bonnard, J.-B. Caillau, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Résumé

Résultats d'optimalité en transfert orbital. Cette Note présente des résultats d'optimalité en transfert orbital. La moyennation du problème de la minimisation de l'énergie est considérée, et les propriétés de la métrique riemannienne associée sont présentées. ***Pour citer cet article : B. Bonnard, J.-B. Caillau, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Version française abrégée

Nous considérons l'équation de Kepler contrôlée (1) qui décrit la dynamique d'un engin spatial dont on commande la poussée. Il s'agit, en transfert orbital, d'amener cet engin d'une orbite elliptique initiale vers une orbite elliptique finale.

Une étape conceptuelle importante consiste à proposer une approximation suffisamment simple du problème pour permettre une étude mathématique complète, notamment des questions d'optimalité. Dans le cas dit *à poussée faible* où la puissance de l'engin est d'amplitude limitée, de nombreuses révolutions sont requises pour réaliser le transfert, et un petit paramètre naturel du système est l'inverse de la longueur angulaire de la trajectoire ou *longitude cumulée* finale. C'est pourquoi après avoir remplacé le coût L^1 associé à la minimisation de la consommation par un coût L^2 (critère *énergie*), une approximation par moyennation est proposée. Le nouvel hamiltonien obtenu s'interprète alors comme celui d'un problème riemannien en dimension trois—on se restreint ici au cas coplanaire, suffisant pour traiter les transferts à faible inclinaison—dont on étudie les propriétés.

Le point essentiel de l'analyse de la métrique est la construction de la forme normale (4) qui met en évidence deux propriétés. La première est l'intégrabilité au sens de Liouvile. La deuxième est qu'on peut se restreindre, par homo-

E-mail addresses: Bernard.Bonnard@u-bourgogne.fr (B. Bonnard), caillau@n7.fr (J.-B. Caillau).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2007.07.028

généité, à l'étude de la restriction de la métrique à l'hémisphère nord d'une 2-sphère de révolution où les méridiens sont en particulier associés au transfert vers des orbites circulaires. Cette métrique peut être prolongée analytiquement à la sphère toute entière et immergée dans une famille de métriques à deux paramètres, c et μ, qui sont des invariants contrôlant notamment le nombre de trajectoires périodiques et le rayon d'injectivité. Leurs valeurs en transfert sont $c = \sqrt{2/5}$ et $\mu = 1/\sqrt{5}$, respectivement.

Le premier résultat donne une condition nécessaire d'optimalité pour la métrique prolongée analytiquement en termes de rayon d'injectivité de sa restriction à la sphère. Alors qu'on en déduit l'existence de points de coupure pour cette métrique, le second résultat utilise un calcul précis de ce lieu de coupure pour finalement montrer que ces mêmes points ne sont pas atteignables dans le cas de la métrique initiale associée au transfert orbital.

Théorème 0.1. [4] *Le rayon d'injectivité de la métrique restreinte à la sphère est égal à* $\mu\pi$*, et une condition nécessaire d'optimalité globale pour la métrique en dimension trois est* $c/\mu \leqslant 1$*. Cette condition n'est pas remplie en transfert orbital* ($c/\mu = \sqrt{2} > 1$) *de sorte qu'il existe des points de coupure pour la métrique prolongée analytiquement.*

Théorème 0.2. [4] *Le lieu de coupure de la métrique restreinte à la sphère est une branche simple incluse dans le parallèle antipodal du point considéré. Les lieux conjugués et les lignes de séparation de la métrique initiale* (*non prolongée*) *en transfert orbital sont par conséquent toujours vides, et seuls des problèmes d'existence dus à l'incomplétude de la métrique peuvent être à l'origine de perte d'optimalité des géodésiques.*

1. Introduction

We consider the control of the Kepler equation that we normalize to

$$\ddot{q} = -q/|q|^3 + u/m \tag{1}$$

where q is the position vector in $\mathbf{R}^n$ ($n = 2$ or 3, depending on whether the coplanar case is considered or not), m the mass of the spacecraft, and u the control, that is the thrust of the engine. The variation of the mass is modelled according to

$$\dot{m} = -\beta|u|, \tag{2}$$

the constant β being related to the physical characteristics of the engine. The control is usually bounded, $|u| \leqslant T_{\max}$, and an important challenge in modern space missions is to use very low thrusts, so that $T_{\max}$ is a small parameter.

An orbit transfer then consists in departing from an initial orbit of the free motion which is periodic when restricted to the so-called *elliptic domain*, and reaching a terminal orbit—for instance a geostationary one in the case of transfers around the Earth. Using the *angular momentum*, $c = q \times \dot{q}$, and the *mechanical energy*, $E = \dot{q}^2/2 - 1/|q|$, the elliptic domain, Q, is defined as

$$Q = \{(q, \dot{q}) \mid c \neq 0,\ E < 0\}.$$

One of the most important performance indexes is the minimization of the fuel consumption. According to Eq. (2), it amounts to minimizing the L^1 norm of the control (the final time, t_f, being fixed):

$$\int_0^{t_f} |u|\,\mathrm{d}t \to \min.$$

The extremal flow of this optimal control problem is described by the Pontryagin maximum principle and stratified by the Lie algebraic structure of the vector fields parameterized by the control. Because of its complexity, numerical simulations are required to compute it [12]. These simulations, however, demonstrate that for low thrusts, only the averaged behaviour of the orbital elements [19] (that is of the geometric coordinates that define the shape and the position of the osculating elliptic orbit) is observed, emphasizing the fibered structure of the elliptic domain. Namely, the fiber is $\mathbf{S}^1$ and Q is equal to $X_e \times \mathbf{S}^1$ where X_e is the *space of ellipses*, defined in terms of first integrals of the unperturbed motion. In the coplanar case where the osculating orbit remains in the same plane, the space of ellipses is for instance defined using the *mean motion*, n, which is a negative power of the semi-major axis of the osculating ellipse ($n = a^{-3/2}$), the *eccentricity* of the ellipse, e, and the *argument of perigee*, θ, which directs the semi-major axis. In these coordinates,

$$X_e = \{n > 0,\ e < 1,\ \theta \in \mathbf{S}^1\}$$

is homeomorphic to the direct product of $\mathbf{R}_+^*$ with the open unit ball of the plane. The coordinate in the fiber is the *longitude*, l, which is an angle determining the position of the spacecraft on its orbit.

In order to provide a complete mathematical analysis of the problem, we consider averaging with respect to longitude since, for low thrust, the control acts as a perturbation of the Kepler equation. Moreover, averaging is performed on a simplified optimal control problem so as to provide tractable computations. We first replace the L^1 cost by an L^2 one (the so-called *energy* criterion), and do not take into account the variation of the mass (2). Though the two criterions are distinct, they are close enough to be connected using a continuation procedure, see [12]. We also restrict ourselves to the coplanar case which is sufficient to analyze transfers with small inclinations. A final relaxation is done, consisting in dropping the bound on the control. The small parameter is then the inverse of the final *cumulated longitude* (counting the number of revolutions required to achieve the transfer), $1/l_f$. The averaged Hamiltonian of the energy minimization is equal to [10]

$$H = (1/2)\left[9n^{1/3}p_n^2 + 5\left(1-e^2\right)p_e^2/\left(2n^{5/3}\right) + \left(5-4e^2\right)p_\theta^2/\left(2n^{5/3}e^2\right)\right].$$

It is quadratic in the adjoint variable $p = (p_n, p_e, p_\theta)$ and of full rank, so that it can be interpreted as the Hamiltonian associated with a three dimensional Riemannian metric on X_e (see [3]),

$$g = \mathrm{d}n^2/\left(9n^{1/3}\right) + 2n^{5/3}\,\mathrm{d}e^2/\left(5\left(1-e^2\right)\right) + 2n^{5/3}e^2\,\mathrm{d}\theta^2/\left(5-4e^2\right). \tag{3}$$

Such an approximation by averaging is $\mathcal{C}^0$-close to the original system and has to be connected to the initial problem using a continuation procedure. The regularity of the path of the corresponding homotopy is besides connected with conjugate point conditions in control theory [6].

We first recall in §2 some geometric preliminaries, stating the existence of a normal form for an analytic extension of the metric. We prove integrability in the Liouville sense and deduce from it that the study of the metric in dimension three can be reduced by projection to studying its restriction to the two-dimensional sphere. We use this fact in §3 to analyze the optimality properties of the metric, and give two results: a necessary condition in terms of injectivity radius on the sphere, and the statement that geodesics of (3) in the space of ellipses may only lose optimality because of existence issues.

2. Geometric preliminaries

A remarkable feature of the geometric coordinates (n, e, θ) is that they are orthogonal [1] for the metric g on X_e. A similar result holds when averaging the energy minimization Hamiltonian of a single-input orbit transfer where the direction of the thrust is fixed and defined by the velocity [5] (the control is thus scalar and reduces to the modulus of the thrust in this prescribed direction).

So as to obtain a normal form of the metric, we set $n = (5r/2)^{6/5}$, $e = \sin\varphi$, and get

$$g = \mathrm{d}r^2 + \left(r^2/c^2\right)\left(\mathrm{d}\theta^2/G(\varphi) + \mathrm{d}\varphi^2\right) \tag{4}$$

which extends analytically the metric from X_e to $X = \mathbf{R}_+^* \times \mathbf{S}^2$, θ and φ being the two standard angles on the sphere. More precisely, $e = \sin\varphi$ and φ within $]0, \pi/2[$ in the space of ellipses X_e, so that the extension replaces one open hemisphere by the whole compact sphere. The second hemisphere is associated to orbits with retrograde orientation, the equator corresponding to straight lines (ellipses with null semi-minor axis) on which both orientations are identical. One has

$$G(\varphi) = \sin^2\varphi/\left(1 - \left(1-\mu^2\right)\sin^2\varphi\right),$$

and the normal form is parameterized by two scalars, c and μ, which in orbit transfer are $c = \sqrt{2/5}$ and $\mu = 1/\sqrt{5}$. The singularities at the poles $\varphi = 0$ (π) are only due to the coordinates themselves, and one retrieves the full rank there by a polar blowing-up.

Since θ is cyclic, p_θ is a linear first integral. A key observation is that $H_2 = (1/2)(p_\theta^2/G(\varphi) + p_\varphi^2)$ is another independent function in involution with H. As a result, the system is Liouville integrable. The second Hamiltonian H_2 also corresponds to a Riemannian metric, namely the restriction g_2 of g to the sphere:

$$g_2 = G(\varphi)\,\mathrm{d}\theta^2 + \mathrm{d}\varphi^2.$$

One readily gets

$$(1/2)(r\dot{r})^2 = H - c^2 H_2 + H r^2,$$

which strongly suggests to define $u = r^2$. Then,

$$\dot{u}^2/\big(H - c^2 H_2 + H u\big) = \mathrm{cst}$$

gives evidence that u is a degree two polynomial in time since H_2 is a first integral. Actually, $\ddot{u} = 4H$, whence the quadrature on r, and the whole system is integrable using elementary functions [4]. Using the time change $\mathrm{d}\tau = c^2\,\mathrm{d}t/r^2$, the computation of the extremal flow of H is reduced by projection to the computation of the extremal flow of H_2 on the sphere. We parameterize geodesics on X by arc length setting $H = 1/2$, so the new time τ is related to the time t by the following expression:

$$\tau(t, r_0, p_{r_0}) = c^2\big(\arctan\big(t/(r_0\cos\alpha_0) + \tan\alpha_0\big) - \alpha_0\big)/(r_0\cos\alpha_0)$$

with $p_{r_0} = \sin\alpha_0$ whenever $p_{r_0} \neq \pm 1$, $\tau = 0$ otherwise (θ and φ remain constant on $H = 1/2$ when $p_{r_0} = \pm 1$). This reduction in two dimensions is also crucial to analyze the optimality properties of the metric.

3. Optimality results

Given a point x_0 of the Riemannian manifold (X, g), we recall that the *cut locus*, $\mathrm{Cut}(x_0)$, is defined as the set of points where geodesics issuing from x_0 cease to be global minimizers [8,11,13,14]. The computation of these sets is the ultimate task to understand the properties of the metric associated with the underlying averaged energy minimization orbit transfer problem. The *separating line*, $L(x_0)$, is the set of points where at least two minimizing geodesics intersect. When the metric is complete, the cut locus is the closure of the separating line [2], and one has just to add the first conjugate points: *conjugate points* are critical values of the exponential mapping, and the set of first such points along geodesics issuing from x_0 is the *conjugate locus*, $C(x_0)$.

In our case, the Riemannian manifold (X, g) turns not to be complete because of the existence of separatrices that reach the singularity $r = 0$ in finite time, and cut loci on X may contain points that are neither in the separating line, nor conjugate points. As a consequence, though geodesics on $\mathbf{S}^2 \simeq \{r = \mathrm{cst}\}$ are always lifted up to geodesics on X, one cannot guarantee that the lift of a minimizing geodesic on the sphere remains minimizing on X. We nevertheless get the following characterization [4].

Given x_0 in X, let y_0 denote its projection on the sphere. Then a point belongs to the conjugate locus $C(x_0)$ of x_0 if and only if it projects into the conjugate locus of y_0 intersected with the open metric ball of radius $c\pi$, $C(y_0) \cap B_o(y_0, c\pi)$, where c is one the two scalar parameters in the normal form (4) of g. Indeed, the time t on $H = 1/2$ and the time σ on $H_2 = 1/2$ (which is the time τ, up to a rescaling due to the difference of level sets for H_2) are related by

$$\sigma(t, r_0, p_{r_0}) = c\big(\arctan\big(t/(r_0\cos\alpha_0) + \tan\alpha_0\big) - \alpha_0\big),$$

so that

$$\sup_{|p_{r_0}|<1}\ \sup_{t\geqslant 0} \sigma(t, r_0, p_{r_0}) = c\pi,$$

the bound not being reached by any geodesic.

Much similarly, a point in the separating line $L(x_0)$ of x_0 projects into the intersection of the separating line of y_0 with the same ball, $L(y_0) \cap B_o(y_0, c\pi)$. Essentially, the analysis is hence reduced to the study of the restriction g_2 of the metric to the sphere, which is a compact (and so, complete) manifold.

The first optimality result we provide uses an upper bound of the Gauss curvature, K, of the metric on the sphere. One has

$$K = \big(\mu^2 - 2\big(1-\mu^2\big)\cos^2\varphi\big)/\big(1 - \big(1-\mu^2\big)\sin^2\varphi\big)^2,$$

so the curvature, which is not everywhere positive (depending on μ, K may take negative values around the poles), is always bounded over by $1/\mu^2$ and this bound is optimal (it is reached on the equator, $\varphi = \pi/2$, where the curvature is constant). According to Rauch theorem, first conjugate times on the sphere are thus bounded below by $\mu\pi$. An important notion with respect to optimality is the *injectivity radius*, which is the infimum over the manifold of distances of points to their respective cut loci:

$$i(\mathbf{S}^2) = \inf_{y_0 \in \mathbf{S}^2} d\big(y_0, \mathrm{Cut}(y_0)\big).$$

Since $\mathbf{S}^2$ is compact, this infimum is reached and we know from the general theory [8] that a point realizing the radius is either a conjugate point, or the half of a closed geodesic. Accordingly, it suffices to estimate first conjugate times and lengths of periodic geodesics to estimate the injectivity radius, whence the following statement:

Theorem 3.1. [4] *The injectivity radius of the metric g_2 on the sphere is equal to $\mu\pi$, and a necessary condition for global optimality of the metric g on X is $c/\mu \leqslant 1$. The condition is not fulfilled in orbit transfer ($c/\mu = \sqrt{2} > 1$) and there exist cut points of the metric in X.*

Going back to the original metric defined not on X but on its submanifold X_e, the question is finally to decide whether such cut points belong to the space of ellipses or not. We compute to this end the cut locus of any point on the sphere.

It is known from Poincaré [17] that the cut locus of a point on the sphere is a finite tree whose extremities are singularities of the conjugate locus when the metric is analytic [16]. Here, g_2 is easily seen to be conformal to the restriction of the flat metric to an oblate ellipsoid of revolution with unit semi-major axis and semi-minor axis equal to μ. Indeed,

$$g_2 = G(\varphi)\,\mathrm{d}\theta^2 + \mathrm{d}\varphi^2 = g_1/\big(1 - \big(1-\mu^2\big)\sin^2\varphi\big),$$

where $g_1 = \sin^2\varphi\,\mathrm{d}\theta^2 + (1-(1-\mu^2)\sin^2\varphi)\,\mathrm{d}\varphi^2$ is the restriction of the flat three-dimensional metric to the aforementioned ellipsoid parameterized by $x = \sin\varphi\cos\theta$, $y = \sin\varphi\sin\theta$, $z = \mu\cos\varphi$. Though cut loci of two conformal metrics may completely differ, this observation is actually crucial in our case since, as in the case of the ellipsoid of revolution, cut loci are reduced single branches. Sufficient conditions to ensure this property on general surfaces of revolution [7,15] are the following [9,18]. If the symmetry with respect to the equator is an isometry, if the Gauss curvature is nonconstant and monotone non-decreasing along half meridians from the north pole to the equator, then the cut locus of any point on the sphere is a subarc of the antipodal parallel of the point. In our case, the curvature is monotone non-decreasing for μ in $[1/\sqrt{3}, 1]$, not in orbit transfer where μ is equal to $1/\sqrt{5}$. To prove that the cut

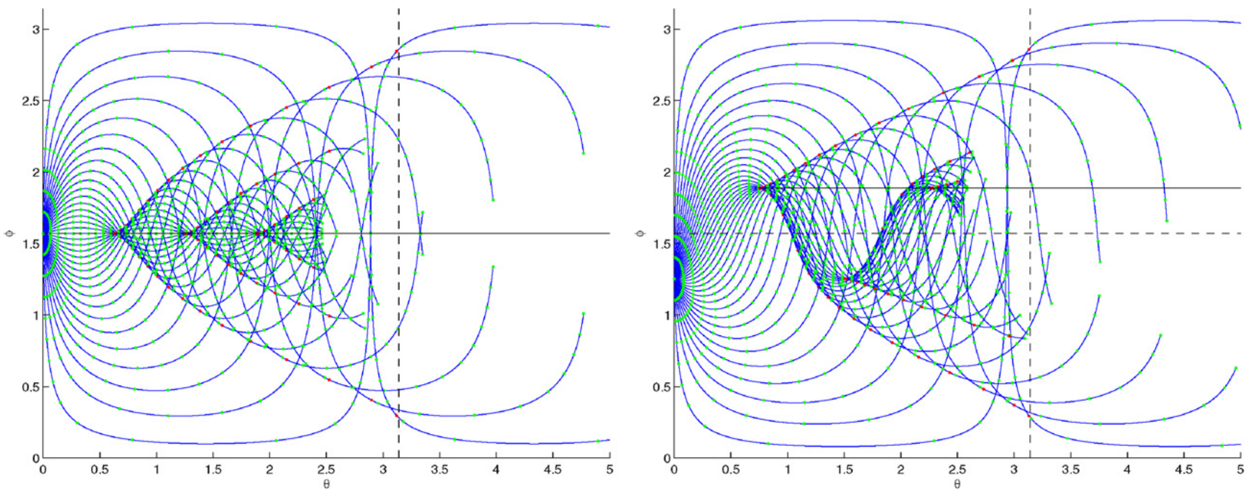

Fig. 1. Conjugate and cut loci of the metric g_2 for $e_0 = 1$ (left) and $e_0 = 9.5e-1$ (right). The first cusps of the conjugate loci corresponding to injectivity radii are clearly observed, while the entire loci have four such cusps forming singularities of astroid-like sets [4]. The cut loci are single arcs contained in antipodal parallels, symmetric to initial points with respect to the equator.

locus still shares the same property, we use estimates of the conjugate points. Indeed, the interior of a domain bounded by two intersecting minimizing curves must contain a conjugate point [16]. Two examples of such conjugate and cut loci computation thanks to the tool developed in [6] are given in Fig. 1.

A direct computation of the separating line (hence of the cut locus, by density) is also possible here using the discrete symmetry group of the metric and the quadratures of the geodesics [4].

According to the analysis of the beginning of the section, conjugate points or points in the separating line in X project onto cut points on the sphere. Such points are antipodal, so they belong to the hemisphere opposite to the initial point and are projections of points in X, not of points in X_e, the space of ellipses.

Theorem 3.2. [4] *The cut locus for the metric g_2 of a point on the sphere is a single branch included in the antipodal parallel of the point. Accordingly, conjugate loci and separating lines of the metric g in orbit transfer on the space of ellipses are always empty, and optimality issues in X_e may only be existence issues due to the lack of completeness.*

References

[1] A. Bolsinov, A. Fomenko, Integrable Geodesic Flows on Two-Dimensional Surfaces, Kluwer, New York, 2000.

[2] R.L. Bishop, Decomposition of cut loci, Proc. Amer. Math. Soc. 65 (1) (1977) 133–136.

[3] B. Bonnard, J.-B. Caillau, Riemannian metric of the averaged energy minimization problem in orbital transfer with low thrust, Ann. Inst. H. Poincaré Anal. Non Linéaire 24 (3) (2007) 395–411.

[4] B. Bonnard, J.-B. Caillau, Geodesic flow of the averaged controlled Kepler equation, HAL preprint number 00134702 (hal.archives-ouvertes.fr/hal-00134702), 2007.

[5] B. Bonnard, J.-B. Caillau, R. Dujol, Energy minimization of single-input orbit transfer by averaging and continuation, Bull. Sci. Math. 130 (8) (2006) 707–719.

[6] B. Bonnard, J.-B. Caillau, E. Trélat, Second order optimality conditions in optimal control with applications, in: Proceedings of the 6th AIMS Conference on Dynamical Systems, Differential Equations and Applications, Poitiers, June 2006, Discrete Contin. Dyn. Syst. Suppl., in press.

[7] G. Darboux, Leçons sur la théorie générale des surfaces, tome III, Gauthiers-Villars, 1914.

[8] M.P. Do Carmo, Riemannian Geometry, Birkhäuser, 1992.

[9] D. Elerath, An improved Topogonov comparison theorem for non-negatively curved manifolds, J. Differential Geom. 15 (1980) 187–216.

[10] R. Epenoy, S. Geffroy, Optimal low-thrust transfers with constraints: generalization of averaging techniques, Acta Astronaut. 41 (3) (1997) 133–149.

[11] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry, Springer, 1990.

[12] J. Gergaud, Th. Haberkorn, Homotopy method for minimum consumption orbit transfer problem, ESAIM Control Optim. Calc. Var. 12 (2006) 294–310.

[13] H. Gluck, D. Singer, Scattering of geodesic field I, Ann. of Math. 108 (1978) 347–372;
H. Gluck, D. Singer, Scattering of geodesic field II, Ann. of Math. 110 (1979) 205–225.

[14] W. Klingenberg, Riemannian Geometry, de Gruyter, 1982.

[15] H. von Mangoldt, Über die jenigen Punkte auf positiv gekrümmten Flächen, welche die Eigenschaft haben, daß die von ihnen ausgehenden geodätischen Linien nie aufhören, kürzeste Linien zu sein, J. Reine Angew. Math. 91 (1881) 23–53.

[16] S.B. Myers, Connections between geometry and topology I, Duke Math. J. 1 (1935) 376–391;
S.B. Myers, Connections between geometry and topology II, Duke Math. J. 2 (1936) 95–102.

[17] H. Poincaré, Sur les lignes géodésiques des surfaces convexes, Trans. Amer. Math. Soc. 5 (1905) 237–274.

[18] R. Sinclair, M. Tanaka, The cut locus of a 2-sphere of revolution and Toponogov's comparison theorem, Preprint, 2006.

[19] O. Zarrouati, Trajectoires Spatiales, CNES–Cepadues, Toulouse, 1987.