C.R. Math. Rep. Acad. Sci. Canada

Vol. XVI,No. 4 **August 1994 août**

CONTENTS

Geometrical integration of the connection equation $X' = XK$ for $so(3)$, $su(2)$, $e(2)$, $sl(2,R)$, and $sl(2,R)^+$

R.V. Chacon F.R.S.C. and A.T. Fomenko

Introduction. The equation of parallel transport, or connection equation, is the equation $X' = KX$ where X is a vector and K the matrix which results when the Christoffel symbols are contracted with the tangent vector. We shall call any equation of this form a connection equation, whether or not K is obtained in this way.

The main goal of the present paper is to present a new method for the direct geometric integration of the connection equation

$$X'(t) = K(t)X(t)$$

for the case when the matrix function $K(t)$ belongs to one of the canonical representations of one of the following three-dimensional classical Lie algebras:

$so(3)$ of the orthogonal group $SO(3)$,

$su(2)$ of the special unitary group $SU(2)$,

$e(2)$ of the group $E(2)$ (motions of Euclidean 2-plane),

$sl(2,R)$ of the group $SL(2,R)$,

$sl(2,R)^+$ of the group $SL(2,R)^+$ (motions of hyperbolic plane),

in the case that the control force is *curvature directed.* It is also shown that the reduction of the general case to this case is equivalent to solving an ordinary differential equation.

Consider the connection equation

$$X'(t) = K(t)X(t)$$

where $a \le t \le b$ is a scalar parameter, $b-a > 0$ is sufficiently small, and the function $K(t)$ takes its value in a fixed matrix Lie algebra g. The solution of the equation can be written in the form.

$$X(t) = e^{H(t)}X(a),$$

where $H(t)$ is a g-valued function on the interval $[a,b]$. In particular, $X(b) = e^{H(b)}X(a)$. Thus, if $K(t)$ is a given curve in the Lie algebra g, then we can correspond to it a new curve $H(t)$ which determines the evolution of solution $X(t)$. Feynman [3] found a formal expression for the solution of the basic equation without proving its convergence. Later, Magnus obtained a well-known differential equation, known as Magnus' equation, for $H(t)$. The problem of calculating $H(t)$ was also considered by Bialynicki-Birula, Mielnik and Plebanski in [4]. In 1976 Karasev and Mosolova [5] obtained an expression for $H(t)$ together with an estimate for its norm, basing their work on the method developed by Maslov. K.T.Chen [6] has obtained some results about $H(t)$ for special finite dimensional algebras. The starting point of the present work is Magnus' equation. There are two forms of the equation, depending on whether right or left multiplication is considered. We shall only make use of the second form.

Magnus' Equation.

a) $H(t)$ in the formula $X(t) = e^{H(t)}X(0)$, where $X(t)$ satisfies the equation $X'(t) = K(t)X(t)$, itself satisfies

$$\frac{dH(t)}{dt} = \frac{\mathrm{ad}_{H(t)}}{e^{\mathrm{ad}_{H(t)}} - 1} K(t) ,$$

and b) . $H(t)$ in the formula $X(t) = X(0)e^{H(t)}$, where $X(t)$ satisfies the equation $X'(t) = X(t)K(t)$, itself satisfies

$$\frac{dH(t)}{dt} = \frac{ad_{H(t)}}{e^{-ad_{H(t)}} - 1}(-K(t)) .$$

Remark. In the right side of the formulas there is an operator function applied to $K(t)$. This function is expressed as an analytic function of the operator $ad_{H(t)}$. As usual it is necessary to consider expansions corresponding to these functions of the form

$$\sum_{i=0}^{\infty} p_i \, ad^i_{H(t)} ,$$

and to apply these operators to $K(t)$ and to $-K(t)$ where

$ad^i_{H(t)}$ is a composition of the operator $ad_{H(t)}$.

Magnus' equations follow at once from the following lemma (see e.g. [3]) since e^H is a non-singular matrix.

Lemma 1. Let $m(z) = (e^z - 1)/z$. Then

$$d(e^{G(t)})/dt = M_1(t)\, e^{G(t)} = e^{G(t)}\, M_2(t),$$

where

$$M_1(t) = m(ad_G)G'(t) \quad and \quad M_2(t) = m(-ad_G)G'(t).$$

Proof:

$$(e^{G(t+c)} - e^{G(t)})/c = (1/c)\int_0^1 \frac{d(e^{sG(t+c)}\, e^{(1-s)G(t)})}{ds}\, ds$$

$$= \int_0^1 e^{sG(t+c)}\left[\frac{G(t+c) - G(t)}{c}\right] e^{(1-s)G(t)}\, ds.$$

As c tends to zero we obtain

$$d(e^{G(t)})/dt = \left[\int_0^1 e^{sG(t)}\, G'(t)\, e^{-sG(t)}\, ds \right] e^{G(t)}.$$

Writing the integrand in a Taylor series about $s = 0$ gives

$$e^{sG(t)}\, G'(t)\, e^{-sG(t)}$$

$$= G'(t) + (ad_G)\, G'(t)\, s + (ad_G)^2\, G'(t)\, (s^2/2) + \dots ,$$

and integrating from zero to one yields part a).

Part b) is obtained in the same way, but starting with the integral of the derivative

with respect to s of $e^{(1-s)G(t+c)}\, e^{sG(t)}$.

Remark. The function $k(z)$

$$k(z) = 1 + \sum_{p=1}^{\infty} k_{2p}\, z^{2p}, \qquad \text{where} \qquad k(z) = \frac{z}{1 - e^{-z}} - \frac{z}{2},$$

used to define Bernoulli's numbers is related to the function $m(z)$

of the lemma: $k(z) = 1/m(z) - (z/2)$.

Notation.

1) R^3 - The commutative Lie algebra of diagonal (3x3) real matrices.
2) $so(3)$ -The Lie algebra of the group $SO(3) = SU(2)/Z_2$ of orthogonal proper rotations of Euclidean 3-space E^3.
3) $su(2)$ - The Lie algebra of the group $SU(2)$ of unitary (2x2)-matrices with unit determinant. The algebras $so(3)$ and $su(2)$
4) are isomorphic.
4) $e(2)$ - The Lie algebra of the group $E(2)$ of isometric proper

transformations of the Euclidean 2-plane E^2.

5) *sl(2,R)* - The Lie algebra of the Lie group *SL(2,R)* of all real (2x2) matrices with unit determinant, so that *SL(2,R)* is the group of symplectic linear transformations of the 2-plane, preserving the area of the plane and having the origin as a fixed point.

6) $sl(2,R)^+$ -The Lie algebra of the Lie group $SL(2,R)/Z_2$ which is the group of proper motions of the hyperbolic plane. In other words, the group $SL(2,R)^+$ is the group of isometries of the 2-dimensional pseudo-sphere in pseudo-Euclidean 3-space of index *1*. The Lie algebras *sl(2,R)* and $sl(2,R)^+$ are isomorphic.

7) $dX(t)/dt = X(t)K(t)$ - the basic differential equation, where $X(t)$ is a vector function and $K(t)$ is a matrix-valued function whose elements are in *g* where *g* is a matrix representation of one of the of Lie algebras mentioned above.

8) The matrix function $K(t)$ can be viewed as a smooth curve in Euclidean space E^3, where this space is identified with one of the Lie algebras R^3, *so(3)*, *su(2)*, *e(2)*, *sl(2,R)*, or $sl(2,R)^+$. The function $K(t)$ is called *the control force*. The norm of E^3 defines a natural norm of *g*, denoted by $|\ |$.

9) It is assumed that the control force *K* may be written in the form $K(t) = \omega(t)k(t)$ where both ω and k are smooth functions of t, where $\omega(t) = \pm|K(t)|$ and where $k(t) = K(t)/\pm|K(t)|$ if $K(t) \neq 0$.

10) *s* denotes arc length along the curve $k(s)$ considered as a curve in the 2-sphere (respectively, the 2-plane, or the hyperbolic plane).

11) The function $H(t)$ is written in the form $H(t) = \theta(t)h(t)$ where $X(t) = X(0)e^{H(t)}$, $H(t)$ belongs to the Lie algebra *g*, and where $|h(t)| = 1$, $\theta = |H|$.

12) $$dH/dt = \frac{ad_H}{e^{-ad_H} - 1}(-K)$$ - Magnus' equation for the equation $X' = XK$.

13) $K = K_{\parallel} + K_{\perp}$ - decomposition of the control force K into the sum of two orthogonal components in E^3. The vector K passes through the point $h(t)$ on the unit 2-sphere S^2, $K_{\parallel}$ is the projection

of K onto the direction of h and $K_\perp$ is the projection of K onto the plane tangent to the sphere S^2 at the point h.

1. Control forces for so(3).

We turn our attention now to the case of the Lie algebra *so(3)*. The case of *su(2)* will be considered separately. The Lie algebra *so(3)* can be identified canonically with skew symmetric real 3 by 3 matrices, and these matrices can be identified in turn with vectors in E^3. These are the identifications used in the discussion of *so(3)*.

Consider the differential equation in E^3:

$$dX(u)/du = X(u)K(u), \qquad (1)$$

$0 \le u \le t$, where $K(u)$ is a *smooth* vector-function with values in *so(3)*. The Lie algebra *so(3)* has been identified canonically with vectors in E^3 and with skew symmetric matrices and thus $X(u)K(u)$ is defined in a natural way as the action of the matrix $K(u)$ on the vector $X(u)$ by right multiplication and is again a 3-dimensional vector.

The solution of (1) may be written in the form $X(t) = X(0)e^{H(t)}$, where $H(t)$ satisfies the second of Magnus' formulas. For example, for the case of the commutative Lie algebra R^n the solution has the form

$$X(t) = X(0)\, e^{\int_0^t K(u)du}$$

so that $H(t) = \int_0^t K(u)\, du$ and in the general case $H(t)$ satisfies

$$H' = \frac{\mathrm{ad}_H}{e^{-\mathrm{ad}_H} - 1}(-K) \quad , \qquad (2)$$

where $\mathrm{ad}_A(B) = [\, A, B\,]$; $\mathrm{ad} : so(3) \longrightarrow so(3)$ is a linear operator.

We note that it is simply a matter of choice whether right or left multiplication is used in the fundamental equation, although it is, of course, necessary to use the formulas which correspond to the choice made. They differ by certain changes in sign which take into account the order in the product integral yielding the solution. This order is opposite for these two cases.

Note that if a smooth curve $K(t)$ is given, the curve $H(t)$ is determined uniquely near the unit. The converse is implied by $K(t) = (e^{H(t)})'$.

Any matrix (vector) P in $so(3)$ can be written in the form $P = \lambda p$, where λ is a scalar and p is a vector. The matrix e^P is 117
identified with the rotation of E^3 having angle of rotation λ about the axis determined by p The matrices in $so(3)$ have a natural norm $|\ \ |$ which coincides with the standard Euclidean norm in E^3. With respect to this norm, $\lambda = \pm |P|$, and if $P \neq 0$, $p = P/|P|$. If P depends on the parameter t then the vector $p(t)$ traces a trajectory on the sphere S^2 of unit radius contained in E^3. We write the control force in terms of this decomposition, $K(t) = \omega(t)k(t)$, and assume throughout that both $\omega(t)$ and $k(t)$ may be chosen to be smooth functions of t.
118 The function $H(t)$ is written in this way also, $H(t) = \theta(t)h(t)$. It follows that $\theta = \pm|H|$, that $\omega = \pm|K|$. It also follows that both $h(t)$ and $k(t)$ are unit vectors at points where the corresponding functions are different from zero. As the parameter t varies the curves $h(t)$ and $k(t)$ trace trajectories on the unit 2-sphere S^2, where $S^2 \subset E^3$.

It follows from its definition that the curve $H(t)$ starts from the point O in 3-space. The *natural parameter s* is defined to be arc-length for the the curve $k(t) \subset S^3$, and it is assumed that each parameter can be solved in terms of the other. The parameter s is taken to be positive in the direction of increasing t, and the point $k_{t=0}$ is taken to correspond to the value $s = 0$.

The authors thank A.Anisov for useful discussions.

2. *Different types of control forces. Reduction to the curvature directed case.*

Definition 1. The differential equation $X' = XK$ and the corresponding control force $K(t)$ will be called *degenerate* on an interval $0 \le t \le T$ if $K(t) = K_0 = const$. In this case the curve $k(t)$ is a point, i.e. a degenerate curve, $k(t) = k_0$ for all t, $0 \le t \le T$.

Definition 2. The differential equation $X' = XK$ and the corresponding control force $K(t)$ will be called *linear geodesic* on an interval $0 \le t \le T$ if $K(s) = \omega_0\, k(s)$, where $\omega_0 \ne 0$ is some constant and the curve $k(s)$ is a segment of a geodesic line on the 2-sphere, i.e. a part of a great circle, and consequently its curvature $r(s)$ is equal identically to zero. Here $r(s)$ is the curvature of the curve on the 2-sphere (not in the ambient E^3). Thus, in the linear geodesic case the point $k(s)$ moves with constant velocity along the geodesic line on the sphere and the norm of $K(s)$ is constant.

Definition 3. The differential equation $X' = XK$ and the corresponding control force $K(t)$ is called *λ-curvature directed* if

$$\lambda\, r(t) = \omega(t) \frac{dt(s)}{ds},$$

where $r(t)$ is the curvature of the curve $k(t)$, $\lambda > 0$ is a *constant* real number and $t(s)$ is the parameter as a function of the arc-length s. In addition the direction of rotation induced is assumed to agree with the direction of rotation of the tangent line to the curve.

If $t = s$, then the λ-curvature directed force $K(s)$ satisfies the following condition:

$$\lambda\, r(s) = \omega(s).$$

Remark. The vector $K(t)$ may be interpreted in the following way: $K(t)\,\Delta t$ represents the rotation of 3-space around the axis $k(t)$ with angle $\omega(t)\,\Delta t$.

In the case of a force which is λ-curvature directed we have:

$$K(t) = \lambda\, r(t) \frac{ds(t)}{dt} k(t),$$

where $r(t)$ is the curvature of the smooth curve $k(t)$ in the 2-sphere (not in the ambient 3-space) and $s = s(t)$ is arc-length. It is assumed also that the curve $k(t)$ is regular with non-zero velocity vector k'. The norm of $k'(s)$ is equal to 1.

Lemma 1. The equation $X' = XK = X\omega k$ is λ-curvature directed if and only if a) after its parameter is changed to arc-length s (along the curve k with a curvature r on the 2-sphere) the norm $\omega(s)$ of the force $K(s)$ is equal to $\lambda\ r(s)$, where $r(s)$ is the curvature of the curve $k(s)$ on the 2-sphere, and b) the direction of rotation induced by K is the same as the direction of rotation of the tangent vector to $k(s)$.

The proof follows at once from the definition of a λ-curvature directed control force: if $X' = XK$, then $X'_s = X\omega t'_s\ k = = X\lambda rk$.

Proposition.

a) It is possible to find a parameter-dependent change of coordinates $X(t) = Y(t)A(t)$ which transforms an arbitrary connection equation into one which is λ-curvature directed, for any $\lambda \neq 0$.

b) This parameter-dependent change $A(t)$ is determined by an scalar ordinary differential equation of third order.

Remark. It is easy to see that there always exists a parameter-dependent coordinate change which transforms the basic equation to one of form $Y' = 0$. We need only consider $X(t) = Y(t)e^{H(t)}$. But to find this transformation we need to solve a problem which is equivalent to the initial problem. The Proposition states that we can find the necessary change of coordinates by solving a specific *ordinary* differential equation of third order, and thus reduce the problem of integrating the basic equation to the curvature-directed case. This last case is shown be solved by a simple geometrical procedure.

Proof. To see this, we substitute $X = AY$ in the equation $X' = KX$ to obtain $AY' = KAY - A'Y$, from which it follows that $Y' = GY$, for a new control force $G = A^{-1}(KA - A')$. It remains to show that A may be chosen in such a way as to make the control force G λ-curvature directed. That condition may be written

$$G(t) = \lambda\ r(t)\ (ds/dt)\ g(t),$$

where $g(t) = G(t)/|G(t)|$, and $r(t)$ is the curvature of $g(t)$. Writing $r(t)$ in terms of $g'(t)$ in this equation yields that the condition for the control force G to be λ-curvature directed is a third order equation for the coefficients of the matrix $A(t)$

The Proposition is proved.

3. *Geometrically similar differential equations.*

Definition 4. Two differential equations

$$X'_t = X(t)K(t) \quad \text{and} \quad Y'_\tau = Y(\tau)P(\tau) \tag{3}$$

will be called *geometrically similar* if after the parameter of each equation is changed to their corresponding arc-lengths we obtain the equations: $X' = X(s)\hat{K}(s)$ and $Y' = Y(s)\hat{P}(s)$ with the following relation:

$$\hat{K}(s) = m(s)\hat{P}(s), \tag{4}$$

where $m(s)$ is a smooth scalar function of s

Remark. It is clear that $k(s) = p(s)$ for geometrically similar equations, i.e. the corresponding normed curves on the 2-sphere coincide.

The notion of geometrically equivalent equations can be reformulated in the following way.

Definition 5. Two differential equations $X'_t = X(t)K(t)$ and $Y'_\tau = Y(\tau)P(\tau)$ are called *geometrically similar* if there exists a time-transformation $t = t(\tau)$ such that

$$K(t(\tau))t'_\tau = m(\tau)P(\tau)$$

for a scalar function $m(\tau)$.

Lemma 2. Definitions 4 and 5 are equivalent.

Proof. We have: $X'_t = X_\tau\ \tau'_t = X(t(\tau))K(t(\tau))$ and $X'_\tau = X(t(\tau))(K(t(\tau)t'_\tau)$. Let us put $X(t(\tau)) = Y(\tau)$, then we obtain: $Y'_\tau = Y(\tau)K(t(\tau))t'_\tau$ and $Y'_\tau = Y(\tau)P(\tau)$. If $K(t(\tau))t'_\tau = m(\tau)P(\tau)$, then with respect to the natural parameter s this collinearity relation will be preserved.The lemma is proved.

If we fix some value of t, then at any point X on S^2 we obtain the vector $XK(t)$ which is tangent to S^2 because $K(t)$ is a skew-symmetric matrix. The vector field XK necessary has zeros on the 2-sphere (when the vector X belongs to the kernel of the operator K). For each non-zero vector XK we can consider the straight line in the direction of this vector, and passing through X (if $XK = 0$, then such a line cannot be determined). Thus, we obtain a 1-distribution $v(t)$ of straight lines tangent to the sphere. This distribution is defined, in general, at all points of the sphere except the two opposite points which correspond to the (in general) one-dimensional kernel of $K(t)$. The distribution $v(t)$ depends on t. As t changes the distribution also changes.

Lemma 3. Two equations (3) are geometrically similar if and only if their 1-distributions are equal for all t..

Let us note that the coefficient of collinearity $m(t)$ depends only on the parameter t and does not depend on the choice of the point $X(t)$ on the 2-sphere.

The proof follows at once.

Remark. The integral curves of two geometrically similar equations can clearly be different.

4. *Every non-degenerate equation $X' = XK$ is geometrically similar to a λ-curvature directed equation.*

Statement 1. Suppose that the normed curve $k(t) = K(t)/\omega(t)$ has non-zero curvature for all values of t in some time-segment. Then the equation $X' = XK$ is geometrically similar in that time segment to a λ-curvature directed equation for any given fixed λ.

Proof. Rewrite the equation $X' = X(t)K(t)$ with respect to the natural parameter s. We obtain $X' = X(s)\hat{K}(s)$, where $\hat{K}(s) = K(t(s))t'$. Consider the new equation

$$Y'_s = Y(s)\,\frac{\hat{K}(s)}{|\hat{K}(s)|}\,\lambda r(s), \tag{5}$$

where $r(s)$ is the curvature of the curve $\hat{k}(s) = k(s)$ on the

2-sphere. Then $m(s) = r(s)/\hat{K}(s)$ and these two equations are geometrically similar. The second equation is clearly λ-curvature directed, so that the statement is proved.

5. *Bicycle wheel coordinate systems and their corresponding rotations.*

Consider the basic equation (1) with respect to an arbitrary parameter and define a moving coordinate system on the 2-sphere in the following way. Let $l(0)$ be the fixed axis in E^3 which is determined by the vector $k(0)$. This axis meets the unit 2-sphere at the point $k(0) = N =$ the north pole.
Consider on the sphere a fixed system of meridians and parallels with north pole N. Define the first meridian or *initial* Greenwich meridian $\mu(0)$ to be that meridian tangent to the curve $k(t)$ at the point $k(0)$, with positive direction in the direction of increasing t. This fixes a bicycle wheel with axis $l(0)$ determined by $k(0)$. Now we start to rotate this coordinate system or bicycle wheel around the axis $l(0)$ with angular velocity $\omega(t)/\lambda$.

Then a second independent motion is imparted to the coordinate system or bicycle wheel: the axis l moves along the curve $k(t)$ on the sphere. More precisely, at time t the axis $l(t)$ has position determined by $k(t)$.

Finally, we can consider the motion of the coordinate system (with marked first Greenwich meridian) which is the composition of these two independent motions:

1) the motion of $l(t)$ along the curve $k(t)$,
2) the rotation of the meridian system with north pole at the point $k(t)$ around the the axis $l(t)$ determined by $k(t)$ with angular velocity $\omega(t)/\lambda$.

This determines a rotating coordinate system on the 2-sphere uniquely.

Definition 6. We call this coordinate system a *(ω/λ)-bicycle wheel system*. For each moment t we obtain a resulting rotation $g(t)_\lambda$ of the 2-sphere. We call this rotation *the resulting (ω/λ)-bicycle wheel rotation.*

This construction may be described in slightly different terms.

Consider the curve $k(t)$, $0 \le t \le T$, and divide it into the union of small segments by considering the set of close points $k(t_i)$. If the north pole is located at the previous step at the point $k(t_{i-1})$ with the Greenwich meridian at a certain location, then the north pole and meridian are moved by parallel transport from point $k(t_{i-1})$ to the point $k(t_i)$ and afterwards the sphere is rotated about the axis $k(t_i)$ by angle $\omega(t_i)(t_i - t_{i-1})/\lambda$. The rotation $g(t)_\lambda$ is the result of applying these steps as $i = 1,\dots,n$, and then taking the limit as the maximal distance between the points tends to zero.

Remark.The (ω/λ)-bicycle wheel coordinate system was defined by a *direct construction*. Consider the mechanical motion of the rigid ball which is induced by the motion of the described coordinate system on the sphere. At each moment t the rigid ball has a vector $D(t)$ as its vector of instantaneous angular velocity. *The vector $D(t)$ in general does not coincide with the vector $K(t)$* The vector $X(0)g(t)_\lambda$ is the solution of the differential equation

$$X' = XD(t)_\lambda.$$

Consider the (ω/λ)-bicycle wheel rotating system at the moment $t = 0$ and the position of this system at some other moment t. The last system is obtained from the initial one by the rotation $g(t)_\lambda$ of 3-space. Consider the position $\mu(t)_\lambda$ of the Greenwich meridian at the moment t. It is clear that the Greenwich meridian $\mu(t)_\lambda$ is the image of the initial Greenwich meridian $\mu(0)$ under the action of the rotation $g(t)_\lambda$. Sometimes we will omit the subscript λ in the notation for the Greenwich meridian.

Consider the following natural questions:

1) How is the rotation $g(t)_\lambda$ calculated as a function of the control force $K(t)$?

2) What is the connection between the axis $k(t)$ and the axis $d(t)_\lambda$?

These questions are addressed in the next sections.

6. *The Main Theorem for so(3).*

The principal result proved in this section is that *if $K(t)$ is 2-curvature-directed, then the equation $X' = XK$ may be solved by a simple geometrical procedure.*

Statement 2. For any smooth control force $K(t)$ the curves $k(t)$ and $h(t)$ start from the same point $k(0) = h(0)$ on the 2-sphere and are tangent at this point.

Theorem 1. Geometrical solution of the equation $X' = XK$ in the degenerate case and in the linear geodesic case.

1) If $K(t) = K_0$ = a constant vector, the degenerate case, then $H(t) = tK_0$, i.e. the vector $H(t)$ is always collinear with K. Thus,

$$X(t) = X(0)e^{tK_0} \tag{6}$$

for all t.

2) When $K(s) = \omega_0 k(s)$ is a linear geodesic control force, i.e. $k(s)$ moves along a great circle with unit velocity $|k'| = 1$ (here $t = s$ is the natural parameter) and with constant norm ω_0, then there exist parameter dependent linear coordinate transformations which can be written explicitly and which reduce the initial equation $X' = XK$ to the degenerate case (1).

Theorem 2. Geometrical solution of the equation $X' = XK$ in the case of a 2-curvature directed control force K.

a) Consider a 2-curvature-directed control force $K(t)$ (i.e. with a curvature $r(t) \neq 0$ and $2r = \omega\, dt/ds$) and a corresponding normed curve $k(t) = K(t)/\omega(t)$ on the 2-sphere. Consider two meridians (great circles) $\mu(0)$ and $\mu(t)$ tangent to the curve k at the starting point $k(0)$ and at the end point $k(t)$ respectively. Then the point of their intersection is exactly the point $h(t)$. The angle between $\mu(0)$ and $\mu(t)$ is equal to $\theta(t)/2$.

b) Thus, the solution of the problem has the form

$$X(t) = X(0)\, e^{\theta(t)h(t)} \tag{7}$$

where $\theta(t)$ and $h(t)$ are described above.

Remarks.

1) The Greenwich meridian $\mu(t)_{\lambda=1/2}$ is always tangent to the *curve k at the point k(t) for any t.* Consequently, the $\mu(t)$ which is tangent to the curve k at the point $k(t)$ *is exactly* the meridian $\mu(t)_{\lambda=1/2}$ for a $(\lambda/2)$-bicycle wheel coordinate system along the curve k. We will denote the meridians $\mu(t)_{\lambda=1/2}$ simply by $\mu(t)$.

2) The evolution of the point $h(t)$ and the value $\theta(t)$ is totally determined by the two tangent great circles to the given curve k an its starting and end points. The first circle is fixed, and the second circle is free to move.

Corollary 1. In the case of a 2-curvature-directed control force K the point h always moves along a fixed great circle on the sphere. This circle is the meridian which is tangent to the curve k at its starting point $k(0)$.

Thus, the point $h(t)$ is the intersection of a fixed circle $\mu(0)$ with the moving circle $\mu(t)$.

Corollary 2. Assume that the curve k has non-zero curvature near its starting point $k(0)$. Then the initial segment of the curve $h(t)$ goes in the direction of convexity of the curve k (outside of k).

7. Proofs for so(3).

7.1. Proof of Theorem 1.

It is easy to see that in the degenerate case, where $K = K_0$, a fixed vector in E^3, that the solution of (1) has the form $X(t) = X(0)e^{K_0 t}$. The curve $k(t)$ then consists of a single point $k = K_0/|K_0|$ (we assume that $K_0 \neq 0$). Then $H(t) = K_0 t$ and is collinear with K , so that $\theta(t) = |K_0|\, t$. Part 1 of Theorem 1 is thus proved.

For the proof of part 2, assume without loss of generality that $\omega_0 = 1$. In that case $X'(s) = X(s)k(s)$, i.e. $\omega(s) \equiv 1$ and $k(s)$ traces a segment of a great circle on the 2-sphere. Because the point $k(s)$ moves along this segment with velocity which has constant norm ($= 1$), the vector $k(s)$ rotates uniformly around a fixed vector *a* in E^3. It follows that if we identify the vector $k(s)$ with its corresponding skew-symmetric matrix, also denoted by $k(s)$, then:

$$k(s) = e^{As} K(s_0) e^{-As}$$

for some constant real skew-symmetric matrix A. Denoting $k(s_0)$ by P_0, we have:

$$k = e^{As} P_0 e^{-As}, \tag{8}$$

where e^{As} is an element of the orthogonal group $SO(3)$. Formula (8) represents the rotation with uniform velocity of the vector P around the fixed axis a. Then we have:

$X'_s = X e^{As} P_0 e^{-As}$ and $X' e^{As} = X e^{As} P_0$, and therefore

$$(X e^{As})' - X (e^{As})' = X e^{As} P_0. \tag{9}$$

Since $\Delta = const$, it follows that $(e^{As})' = A e^{As} = e^{As} A$.
We use here the fact that the matrices e^{As} and A commute for constant A.
It follows from (9) that:

$$(X e^{As})' = X e^{As} A + X e^{As} P_0 = (X e^{As})(A + P_0).$$

Let $Y = X e^{As}$ then $Y' = Y (A + P_0)$, or

$$Y' = Y B_0, \quad B_0 = A + P_0 .$$

Theorem 1 is proved.

7.2. *Proof of Statement 2.*

Consider the smooth curve $k(t)$ starting from the point $k(0)$ and the smooth curve $h(t)$ starting from the point $h(0)$. We want to prove:

The two curves $k(t)$ and $h(t)$ (where $t \geq 0$) start from the same point, i.e. $k(0) = h(0)$ and these curves are tangent at that point. This implies that the initial meridian $\mu(0)$ is tangent to both curves $k(t)$ and $h(t)$ for $t = 0$.

Consider the curve $K(t)$ and a small time interval Δt
Since the function $K(t)$ is assumed to be smooth, it can be

approximated on Δt by a linear segment with ends at $(0,K(0))$ $(\Delta t,K(0)+\Delta K)$. This linear approximation has the form: $\hat{K}(t) = K(0) + t\Delta K/\Delta t$, $0 \le t \le \Delta t$. Let us calculate the approximation of the function $H(\Delta t)$. We have:

$$H(\Delta t) = H_1 - H_2 + \dots ,$$

(alternating signs) where the first two homogeneous terms are given in section 3 of [9] to be

$$H_1 = \int_0^{\Delta t} K(u)\,du, \quad H_2 = (1/2)\int_{u=0}^{\Delta t} [\,K(u), \int_{v=0}^{u} K(v)\,dv\,]du.$$

We may therefore write $H_1 = \int_0^{\Delta t} (K(0) + u\Delta K/\Delta t)du = K(0)\,\Delta t + \Delta t\Delta K/2$ $= (K(0) + \Delta K/2)\Delta t$.

Further, $H_2 = (1/2)\int_{u=0}^{\Delta t} [K(0) + u\Delta K/\Delta t\, , \int_{v=0}^{u} (K(0) + v\Delta K/\Delta t)dv\,]du =$

$$=(1/2)\int_0^{\Delta t} [K(0) + u\Delta K/\Delta t\, ,\ K(0)u + u^2\Delta K/2\Delta t\,]du =$$

$$=(1/2)\int_0^{\Delta t} ([K(0),\ u^2\Delta K/2\Delta t\,] + [\,u\Delta K/\Delta t\, ,\ K(0)u\,] + [\,u\Delta K/\Delta t\, ,$$

$$u^2\Delta K/2\Delta t\,])du = (1/2)\int_0^{\Delta t} [K(0),\ \Delta K\,]\,(\,u^2/2\Delta t - u^2/\Delta t\,)du =$$

$$= -\frac{1}{12}(\Delta t)^2\,[K\, ,\Delta K\,].$$ Finally, we obtain the following formula:

$$H(\Delta t) = (K(0) + \tfrac{1}{2}\,\Delta K\,)\Delta t + \tfrac{1}{12}\,(\Delta t)^2\,[K\, ,\Delta K\,] + \dots$$

Note that $K(0) + (1/2)\,\Delta K$ is approximately equal to $K(\Delta t/2)$. Normalizing all terms, we obtain that the vector $h(\Delta t)$ is approximately equal to the sum of the following vectors:

$$h(0) + k(\Delta t/2) + \varepsilon\,[k(0),\ \Delta k(0)],$$

where ε is a small constant and therefore $h(0) = k(0)$. More precisely, $\varepsilon = (\Delta t)^2/12$ and is of second order with respect to Δt. Statement 2 is proved.

Remark. The vector $[K(0), \Delta K]$ is orthogonal to both vectors $K(0)$ and ΔK, because the commutator $[\ ,\ ]$ is the usual cross-product of vectors in 3-space. Therefore the triangle with vertices $k(0) = h(0)$, $k(\Delta t)$ and $h(\Delta t)$ has segments $k(0)-h(\Delta t)$ and $h(\Delta t)-k(\Delta t)$ of approximately equal length.

7.3. Transition from so(3) to su(2). Simplification of Magnus' equation for the case of the Lie algebra su(2).

In this section a calculation is given of an equation which is equivalent to Magnus' equation in the special case of su(2). First recall some standard facts about *so(3)*. This Lie algebra is the linear space of all 3×3 skew-symmetric matrices X, $X^{tr} = -X$. The Lie algebra *so(3)* is isomorphic to the Lie algebra *su(2)* of the group *SU(2)*. The equation is used for the analysis of *su(2)*, and the isomorphism then yields information about the solution of Magnus' equation for *so(3)*.

The Lie algebra *su(2)* is the linear space of all skew-Hermitian (2x2)-matrices with zero trace. The three matrices

$$\tilde{e}_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \tilde{e}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \tilde{e}_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

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

$$[\ \tilde{e}_1/2,\ \tilde{e}_2/2\ ] = \tilde{e}_3/2$$

....

....

Each matrix $\tilde{H}$ from *su(2)* can be represented in the form $\theta\tilde{h}$, where $\tilde{h}^2 = -E$, E = identity matrix. The matrices of *su(2)* are denoted by a letter with the symbol ~ over the letter in order to distinguish them from the matrices of *so(3)* which are denoted by plain letters.

The three matrices

$$e_1 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

form a basis for so(3). Their commutation are:

$$[\ e_1,\ e_2\ ] = e_3,$$

....

....

Each matrix H from $so(3)$ has the form $H = \theta h$, where θ is the angle of rotation of 3-space around the axis determined by the unit vector h.

To establish the isomorphism between $su(2)$ and $so(3)$ we define the following mapping:

$$V: \ \tilde{e}\ /2 \ {-}{-}{-}{-}{-}{>}\ e\ .$$

The matrix $\theta\tilde{h}/2$ is mapped by V to the matrix θh, so that

$$e^V \ : \ e^{\theta\tilde{h}/2} \ {-}{-}{-}{-}{-}{>}\ \exp\theta h,$$

where the symbol e is used for the exponent for $SU(2)$ and the symbol exp for $SO(3)$.

These formulas are a reflection of the fact that the group $SU(2)$ is a 2-fold covering over the group $SO(3)$. Thus, if we take a matrix $\exp H = \exp\theta h$ in $SO(3)$, where θ is the angle of rotation in E^3, then one of its two preimages in $SU(2)$ is a matrix which has norm $\theta/2$.

In what follows we shall work only in $su(2)$ will omit the symbol ~ for simplicity. To go from $su(2)$ to $so(3)$ we need simply to multiply all norms (in particular, θ and ω) by the coefficient 2.

Thus, we can write any matrix (vector) H from $su(2)$ in the form $H = \theta h$, where $h^2 = -E$ where E is the identity matrix. We can also introduce a norm $|\ \ |$ in $su(2)$ in such a way that $\theta = |H|$ and $|h| = 1$. When t changes we obtain a curve $h(t)$ on the 2-sphere S^2, given a curve $H(t)$ in $su(2)$.

Any element Y from $su(2)$ is a skew-Hermitian complex

(2 x 2)-matrix with zero trace:

$$Y = \begin{pmatrix} ia & z \\ -\bar{z} & -ia \end{pmatrix} = \begin{pmatrix} ia & u + iv \\ -u + iv & -ia \end{pmatrix},$$ where a, u, v are real numbers.

Remark. It is also possible to rewrite our calculations and theorems in the terms of quaternions, suggesting possible extensions to the case of $so(n)$ using Clifford algebras.

Lemma 4. Each matrix Y from the Lie algebra su(2) can be written uniquely in the following form:

$$Y = \theta h , \qquad (13)$$

where $\theta \ge 0$, $\theta = | Y | = (\frac{1}{2} \text{ Trace } Y \bar{Y}^T)^{1/2}$, *and h is a skew-Hermitian (2 x 2)-matrix with zero trace, such that* $h^2 = -E$, *where E is the identity matrix..*

Proof. If $Y \in su(2)$, then $Y^2 = -\theta^2 E$ where E is the 2 x 2 identity matrix and $\theta = (a^2 + u^2 + v^2)^{1/2} = | Y | =$

$= [(1/2) \text{ trace } YY^T]^{1/2}$. This implies that $Y = \theta h$ where $|h| = 1$

(hence $h^2 = - E$) and $\theta = |Y|$. .Since $H = H(t)$ belongs to $su(2)$, we can write $H(t) = \theta(t) h(t)$, where $\theta = | H |$; $h^2 = -E$; $| h | = 1$.

Lemma 5. The following formula for the exponential mapping exp: su(2) ----> SU(2) is valid:

$$e^H = E \cos \theta + h \sin \theta. \qquad (14)$$

The proof follows from the expansion: $e^H = \sum_n \frac{1}{n!} H^n$ and $H = \theta h$, $h^2 = -E$.

If H is a function of t, then so are θ and h. Thus, $h = h(t)$ is represented by a curve on the unit sphere S^2 in 3-space $E^3 = su(2)$. Consider the vector-function $h' = \frac{d}{dt} h(t)$ in E^3 and its norm $| h' |$.

Remark. The norm $|X|$ of the matrix X in $su(2)$ is equal to the usual Euclidean length of the vector X representing the matrix X in $R^3 = su(2)$.

Consider the point $h(t)$ on the sphere S^3 and attach the given known vector $K(t)$ to the point $h(t)$. Consider the tangent plane $T_h S^2$ to the sphere S^2 at the point $h \in S^2$. Then the vector $K(t)$ *can be written as the sum of two vectors:* $K = K_{\parallel} + K_{\perp}$, where $K_{\parallel}$ is parallel to the radius-vector h, and $K_{\perp}$ is in the plane $T_h S^2$. For every t we thus have a *unique* decomposition:

$$K(t) = K_{\parallel} + K_{\perp} \tag{15}$$

at the point $h(t)$.

Statement 4.

1) Magnus' equation

$$K(t) = e^{-H(t)}(e^{H(t)})', \quad \text{or } H' = \frac{ad_H}{e^{-ad_H} - 1}(-K)$$

for the function $H(t)$ in the Lie algebra su(2) is equivalent to the following vector equation in 3-space E^3:

$$\theta' h + h' \cos\theta \sin\theta - hh' \sin^2\theta = K, \tag{16}$$

where hh' denotes the usual product of two (2 x 2)-matrices h and h' from the Lie algebra su(2).

2) The previous equation (16) is also equivalent to the following system of two vector equations in 3-space E^3:

$$\theta' h = K_{\parallel} ,$$

$$|h'| \sin\theta (\alpha \cos\theta + \alpha \times h \sin\theta) = K_{\perp} . \tag{17}$$

Equations (17) are the orthogonal projections of equation (16) on two orthogonal subspaces in E^3: the tangent space to sphere at the point $h(t)$ and on the radius-vector $h(t)$. In equation (17) $\alpha \times h$ is the usual cross-product of two vectors in E^3 and $H(t) = \theta(t)\, h(t)$, $K(t) = K_{\parallel}(t) + K_{\perp}(t)$.

3) Magnus' equation can also be written in the following form (one vector equation and one scalar equation):

$$h' = (h + E \, ctg \, \theta) K_\perp \qquad \theta' = \langle K, h \rangle \tag{18}$$

Here we multiply the (2 x 2)-matrices from su(2) in the usual sense.

Remark. It is easy to see that the product hh' of two matrices h and h' from $su(2)$ (where $h^2 = -E$) again belongs to $su(2)$ (see the proof below).

Proof. Differentiate with respect to t:
$(e^{H(t)})' = (\cos\theta + h \sin\theta)' = - \theta' \sin\theta + h' \sin\theta + h \theta' \cos\theta$. We then have
$K(t) = e^{-H(t)}(- \theta' \sin\theta + h' \sin\theta + h \theta' \cos\theta)$ and thus
$(\cos\theta - h \sin\theta) (- \theta' \sin\theta + h' \sin\theta + h \theta' \cos\theta) = K$.
This implies
$- \theta' \cos\theta \sin\theta + h' \cos\theta \sin\theta + h \theta' \cos^2\theta + h \theta' \sin^2\theta - h h' \sin^2\theta - h^2 \theta' \sin\theta \cos\theta = K$.
Since $h^2 = -E$, this implies that
$h' \cos\theta \sin\theta + h \theta' - h h' \sin^2\theta = K$, which is equation (16). We have proved the first part of Statement 4.

Write $h' = |h'| \alpha$, *where* $\alpha^2 = -E$. Since $h = 1$ it follows that the vector α is orthogonal to the vector h. Since h and h' belong to $su(2)$, they are skew-Hermitian matrices with zero trace.

Since $h^2 = -E$, we have $(h^2)' = 0$ and so $hh' + h'h = 0$, and hence $hh' = - h'h$. Thus,
$hh' = \frac{1}{2} (hh' - h'h) = \frac{1}{2} [h,h'] = h \times h'$, where $h \times h'$ is the usual cross-product of vectors in E^3. In particular (because $hh' = \frac{1}{2} [h,h']$), the product hh' again belongs to $su(2)$.

Remarks. *a)* The equation $e^{H(t)}K(t) = (e^{H(t)})'$ follows at once if $X(t) = X(0) e^{H(t)}$ is substituted into the connection equation $X'(t) = X(t) K(t)$. The substitution yields

$$X(0)e^{H(t)}K(t) = X(0)(e^{H(t)})',$$

from which the equation follows, since it holds for all $X(0)$.

b) A direct calculation (since $a^2 + u^2 + v^2 = 1$) shows that

$$\begin{pmatrix} ia & u+iv \\ -u+iv & -ia \end{pmatrix} \begin{pmatrix} ia' & u'+iv' \\ -u'+iv' & -ia' \end{pmatrix} =$$

$$\begin{pmatrix} \underbrace{-aa' - uu' - vv'}_{\text{zero}} + i(uv' - vu') & (-av' + va') + i(au' - ua') \\ (-va' + av') + i(-ua' + au') & \underbrace{-vv'-uu'-aa'}_{\text{zero}} + i(-uv' + vu') \end{pmatrix}.$$

Thus, equation (16) may be rewritten in the following way:

$$h\ \theta' + h' \sin\theta\ (\ \alpha\cos\theta - h \times \alpha \sin\theta\ ) = K.$$

Since both vectors α and $\alpha \times h$ are orthogonal to h, the vector $\alpha\cos\theta + \alpha \times h \sin\theta$ is tangent to the 2-sphere S^2 at the point $h(t)$. Thus, $h\ \theta' = K_{\parallel}$ and

$$h' \sin\theta\ (\ \alpha\cos\theta + \alpha \times h \sin\theta\ ) = K_{\perp}.$$

The second part of statement 4 is proved.

Equation (16), that

$$h' \cos\theta \sin\theta + h\ \theta' - h\ h' \sin^2\theta = K = K_{\parallel} + K_{\perp}$$

implies

$$\theta' h = K_{\parallel} \quad \text{and} \quad h' \sin\theta\ (\ E\cos\theta + h\sin\theta\ ) = K_{\perp}\ ,$$

and thus

$$h' \sin\theta\ e^{\theta h} = K_{\perp}\ .$$

We therefore have that

$$h' = K_{\perp} \frac{e^{-\theta h}}{\sin\theta} \quad \text{and} \quad \theta' = \langle K, h \rangle\ .$$

The first equation may be written

$$h' = K_\perp \left(\frac{E \cos\theta - h \sin\theta}{\sin\theta} \right) = K_\perp (E \,ctg\, \theta - h).$$

Since $K_\perp$ is orthogonal to h, it follows that $K_\perp h = -h K_\perp$ and thus $h' = (h + E \,ctg\, \theta) K_\perp$. The 3rd part of statement 4 is therefore proved, and this proves all of statement 4.

Remark. Magnus' equation can be also be given in the following equivalent form:

$$\begin{aligned} h' &= \frac{1}{2} [h, K] + K_\perp \,ctg\, \theta \\ \theta' h &= K_\perp \end{aligned} \tag{19}$$

Since h and $K_\perp$ anticommute

$$\frac{1}{2} [h, K] = \frac{1}{2} [h, K_\perp + K_\parallel] = \frac{1}{2} [h, K_\perp] = \frac{1}{2} (h K_\perp - K_\perp h)$$

$= h K_\perp$, and thus (19) holds..

7.5. *Geometrical interpretation of Magnus' equation for su(2).*

The vector function $K(t)$ is the given function, but we know its decomposition $K = K_\parallel + K_\perp$ only if we fix the point $h(t)$ on the 2-sphere. This dependence may be written explicitly : $K(t) = K_\parallel(h(t)) + K_\perp(h(t))$. The left side of this formula is known for all t, but the two vectors $K_\parallel$ and $K_\perp$ are known only if we fix the point $h(t)$ on the 2-sphere.

Consider the tangent plane to the 2-sphere at the point $h(t)$. We have two orthogonal unit vectors α and $\alpha \times h$, where $\alpha = h'/|h'|$, in this plane. The vector $K_\perp$ also belongs to the same plane. Consider the point P on the vector $\alpha \times h$ at distance $|h'|$ from the initial point $h(t)$.

Lemma 6. The second of Magnus' equations (17) (which contains the force $K_\perp$) is equivalent to the following geometrical picture: the end of the vector $K_\perp$ always belongs to the circle S^1 on the tangent plane, tangent to the vector α at the point $h(t)$, *having diameter h', and center along $\alpha \times h$. The angle between the vectors $K_\perp$ and α is equal to θ (in particular, we obtain a geometrical interpretation of the function $\theta = |H|$).*

Proof. Write $h' = |h'|\ \alpha$ and rotate the vector α in the positive direction by angle θ. We obtain the vector $\alpha \cos\theta + \alpha \times h \sin\theta$. If this vector is multiplied by the scalar $h' \sin\theta$, we obtain a point M on the circle S^1, and the angle between the diameter and the segment MP is equal to θ, where P is the point on S^1 determined by $\alpha \times h$. The vectors α and $\alpha \times h$ are unit vectors and are orthogonal, and therefore $\alpha \cos\theta + \alpha \times h \sin\theta$ is a unit vector. By (17) $K_\perp = h' \sin\theta\ (\alpha \cos\theta + \alpha \times h \sin\theta)$, and the lemma is proved.

Lemma 7. If K is the control force, then Magnus' equation gives the following geometrical picture: the angle between h' and $K_\perp$ is equal to θ.

The proof follows immediately from Lemma 4.

Lemma 8. The following formulas are valid:

$$\theta' = |\ K_\parallel\ |, \quad |\ h'\ | = |\ K_\perp\ |\ /\ \sin\theta. \tag{20}$$

The proof follows immediately from (17).

7.6. Decomposition of Magnus' equation in su(2) with respect to meridians and parallels.

The vector $K(t)$ is the axis of the bicycle-wheel coordinate system on the 2-sphere, and is the vector of angular velocity of meridians around the axis $K(t)$ at the moment t. The end of the vector $k(t)$ can be taken to be the north pole $N(t)$. Let μ be the meridian and π the parallel passing through $h(t)$ on S^2. Let L be the vector which is the linear velocity of the point h induced by the rotation of the meridian μ around the axis $k(t)$ with angular velocity ω, and let vector $R = \operatorname{ctg}\theta\ K_\perp$. It follows that the vector L is tangent to the parallel π and that the vector R is tangent to the meridian μ.

Lemma 9. The equation $h' = L + R$ is equivalent to the vector part of Magnus' equation (18), and is the orthogonal decomposition of the velocity vector h' into the sum of two *orthogonal vectors tangent to the sphere in the direction of the parallel π and the meridian μ. Thus Magnus' equation may be written*

$$h' = L + R \text{ and } \theta' = \langle K, h \rangle \qquad (21)$$

Further,

$$L = \omega \sin y, \quad R = \omega \operatorname{ctg} \omega \sin y, \quad \theta' = \omega \cos y \qquad (22),$$

where y is the angle between the vectors h and k.

Proof. *Since* L is the vector of angular velocity of the point of the meridian μ which is occupied at time t by $h(t)$, it is tangent to π and $|L| = \omega \sin y$. Also $|K_{\parallel}| = \omega \cos y$ and $|K_{\perp}| = \omega \sin y$.

The vector $K_{\perp}$ is tangent to the meridian μ because $K_{\perp} = K - K_{\parallel}$. The vector $K_{\perp}$ is always oriented in the direction of the north pole N. Thus, the vector $hK_{\perp}$ is orthogonal to the vectors h and $K_{\perp}$ and consequently, $hK_{\perp}$ is parallel to the vector L. On the other hand, we have that $|hK_{\perp}| = |K_{\perp}| = |K| \sin y = \omega \sin y$, i.e. $|L| = |hK_{\perp}|$, and thus $L = hK_{\perp}$.

Since $R = \operatorname{ctg} \theta \, K_{\perp}$, it follows from (18) and the equality $L = hK_{\perp}$ that $h' = L + \operatorname{ctg} \theta \, K_{\perp} = L + R$. The lemma is proved.

7.6. Proof of Theorem 2 for so(3).

Let us return to *so(3)* and consider a 2-curvature-directed control force K. Then Magnus' equation may be written

$$h'_{\mu} = L \text{ and } h'_{\pi} = R,$$

$$|h'_{\mu}| = |L| = \frac{\omega}{2} \operatorname{ctg} \frac{\theta}{2} \sin y,$$

$$|h'_{\pi}| = |R| = \frac{\omega}{2} \sin y,$$

$\theta'/2 = (\omega/2) \cos y$ or: $\theta' = \omega \cos y$,

the angle between h' and meridian μ is equal to $\theta/2$,

where h'_{μ} is the projection of h' on the meridian and h'_{π} is the projection of h' on the parallel, the direction orthogonal to μ.

Lemma 10.

a) If $K(t)$ is a 2-curvature directed control force in $so(3)$, then the Greenwich meridian $\mu(t)$ is always tangent to the curve k at the point $k(t)$, where the meridian $\mu(t)$ is the image of the initial meridian $\mu(0)$ at the moment t under the rotation of the $(\omega/2)$-bicycle wheel coordinate system.

b) The point $h(t)$ does not move in direction of the parallel with respect to the moving $(\omega/2)$-bicycle wheel coordinate system on the 2-sphere. Thus, the point $h(t)$ always belongs to the moving meridian $\mu(t)$ and can move only along $\mu(t)$ and the point $h(t)$ is exactly the point of intersection of the moving meridian $\mu(t)$ with the curve $h(t)$.

Proof. a) Assume for simplicity that the parameter t along the curve k is arc-length, $t = s$. Then the first statement of the Lemma follows from the well known interpretation of the curvature of the curve as the velocity of rotation of the tangent vector to the curve. The initial Greenwich meridian $\mu(0)$ is tangent to the curve k at the point $s = 0$. The meridian $\mu(s)$ has the form $g(s)\mu(0)$, where $g(s) = g(s)_{1/2}$ is the resulting $(\omega/2)$-bicycle wheel rotation. Because K is 2-curvature-directed, we have that $\omega(s) = 2r(s)$, where $r(s)$ is the curvature of k (in the sphere, not in ambient E^3). It is well known that the derivative of the angle rotation of the velocity vector of the plane curve is equal to the curvature of the curve with respect to arc-length. This concludes the proof of part a) in the case that the parameter is arc-length. The case of arbitrary parameter t obviously follows from this.

b) Consider the moving coordinate system. Because L is the orthogonal projection of the velocity vector h' in the direction of the meridian, we see that the component of the velocity h' in the direction of the parallel coincides with the linear velocity of the point $h(t)$ of the meridian induced by the rotation of the moving coordinate system around the axis $k(t)$. The motion of the point $h(t)$ of the meridian is the sum of two motions: the motion along the tangent line (the velocity of which is equal to k') and the rotation of the meridian about the point $k(t)$ with a velocity $\omega(t)/2$. This means that the horizontal component of the *relative* velocity of $h(t)$ with respect to the moving coordinate system is identically zero. Thus, the point h moves only along the meridian μ in the moving coordinate system. The lemma is proved.

Lemma 11. The point $h(t)$ remains on the initial meridian $\mu(0)$, so that $h(t)$ is the point of intersection of the meridians $\mu(0)$ and $\mu(t)$.

Proof. We will give two proofs. The first proof is based on the properties of the product integral. Recall that

$$g(t) = e^{H(t)} = \lim_{N\to\infty} e^{\frac{t}{N}K_1} \cdots e^{\frac{t}{N}K_N}$$

where $K_i = K((ti)/N)$, $i = 1,\dots,N$.

Consider motion along the curve $k(t)$ in the opposite direction: from the point $k(t)$ to the point $k(0)$. Then the resulting rotation $\bar{g}(t)$ similarly equals

$$\bar{g}(t) = e^{\bar{H}(t)} = \lim_{N\to\infty} e^{\frac{-t}{N}K_N} \cdots e^{\frac{-t}{N}K_1}$$

Repeating the arguments used for the rotation $g(t)$ we obtain that the Greenwich meridian $\bar{\mu}(0) = \mu(t)$ is tangent to the curve $\bar{k}$ at the point $\bar{k}(0) = k(t)$. Further, $\bar{h}(t) \in \bar{\mu}(t) = \mu(0)$, and from

$$e^{\bar{H}(t)} = \lim_{N\to\infty} e^{\frac{-t}{N}K_N} \cdots e^{\frac{-t}{N}K_1} = \lim_{N\to\infty} \left(e^{\frac{t}{N}K_1} \cdots e^{\frac{t}{N}K_N} \right)^{-1} =$$

$$= e^{-H(t)}$$

it follows that $\bar{h}(t) = -h(t)$. This implies that $h(t) \in \mu(0)$. The lemma is proved. We also give a second proof.

Consider the time-interval $(0,t)$ and introduce the parameter u along the curve k from $u = 0$ up to $u = t$. Then we consider another differential equation connected with the initial one, namely:

$$\bar{X}'(v) = \bar{X}(v)K(v),$$

where $v = t - u$, $u = t - v$ and $\bar{X}(v) = X(t - v) = X(u)$.

Then we have:

$$\overline{X}'(v) = \frac{d}{dv} X(t - v) = - \frac{d}{du} X(t - v) = - \frac{d}{du} X(u).$$

Thus, $\overline{X}'(v) = - X'_u(u) = - X(u) K(u) = \overline{X}(v) (- K(u))$.

On the other hand, $\overline{X}'(v) = \overline{X}(v) \overline{K}(v)$. From this comparison we see that

$$K(v) = - K(u).$$

Thus, $\overline{X}(v) = \overline{X}(0) e^{\overline{H}(v)}$, $\quad \overline{X}(t) = \overline{X}(0) e^{\overline{H}(t)}$,

$X(0) = X(t) e^{\overline{H}(t)}$ and $X(t) = X(0) e^{-\overline{H}(t)}$. This implies $H(t) = - \overline{H}(t)$. Then we repeat the arguments of the first proof.

Remark. The proof of Lemma 11 shows that the inverse motion along the curve k corresponds to the force $G(u) = -F(t - u) = -K(u)$. The motion from $k(0)$ to $k(t)$ rotates the tangent meridian μ in a counter-clockwise way. The motion from $k(t)$ to $k(0)$ rotates the tangent meridian in a clock-wise direction.

Theorem 2 is proved.

8. Direct geometrical integration of $X' = XK$ in su(2) for 1-curvature directed forces K.

Theorem 3.

a) Consider the equation $X' = XK$ on su(2), where the control force $K(t)$ is 1-curvature-directed (i.e. with a curvature $r(t) \neq 0$ and $r = \omega\, dt/ds$) and the corresponding normed curve $k(t) = K(t)/\omega(t)$ on the 2-sphere. Consider two meridians (great circles) $\rho(0)$ and $\rho(t)$ tangent to the curve k at the starting point $k(0)$ and at the end point $k(t)$ respectively. Then the point of their intersection is exactly the point $h(t)$. The angle between $\rho(0)$ and $\rho(t)$ is equal to $\theta(t)$.

b) The solution of the problem for su(2) has the form

$$X(t) = X(0) e^{\theta(t)h(t)}$$

where $\theta(t)$ and $h(t)$ are described above.

Remark. In this case it is necessary to consider a ω-bicycle wheel coordinate system, as opposed to the $(\omega/2)$-bicycle wheel coordinate system of $so(3)$. The *Greenwich meridian* $\rho(t)_{\lambda=1}$ is tangent to the curve $k(t)$ for all values of t. The evolution of the point $h(t)$ and the scalar $\theta(t)$ is determined by the two tangent great circles, the first at the starting point, and the second at $k(t)$.

Corollary 3. In the case of a 1-curvature directed control force K on $su(2)$ the point h always moves along a fixed great circle on the sphere. This circle is the meridian which is tangent to the curve k at its starting point $k(0)$.

The proof follows from section 7 by replacing $\theta/2$ and $\omega/2$ by θ and ω in all formulas.

9. Direct geometric integration of the equation $X' = XK$ for $e(2)$ in the 2-curvature directed case.

Consider an infinitely small neighborhood of the point $k(0)$ on the sphere S^2 and assume that we analyze the evolution of the system $X' = XK$ only in this infinitely small region of the sphere. Then the action of the Lie group $SO(3)$ in this region is equivalent to the action of the group $E(2)$ of motions of the Euclidean 2-plane. This implies that Theorem 2 is still valid for $e(2)$-case without any major changes.

We can identify the Lie algebra $e(2)$ with Euclidean 3-space. Fix a 2-plane (in 3-space = $e(2)$) which represents the Euclidean plane R^2 where the Lie group $E(2)$ acts (by rotations and parallel transport). Recall that any element of the Lie group $E(2)$ is a rotation about a point of the 2-plane, or else parallel transport by a certain distance in a certain direction on the 2-plane.

The vector function $K(t)$ will be assumed to be in the part $e^r(2)$ of the algebra $e(2)$ which is mapped to rotations under the exponential map. This part of the algebra can be naturally identified with vectors of length $\omega(t)$ which are orthogonal to R^2, the initial point representing the point about which the rotation takes place.

The curve $k(t)$ of initial points on the plane of vectors of unit length ($= K(t)/\omega$) is a plane curve and its curvature is the usual curvature of the curve in the Euclidean plane. We have thus replaced S^2 of $so(3)$ by the 2-plane. The point $h(t)$ traces a trajectory on the 2-plane which is tangent to the curve k at the starting point.

Theorem 3. Consider the equation $X' = XK$ on the part $e^r(2)$ of the Lie algebra $e(2)$ of the Lie group $E(2)$ of motions of R^2 for the case of a 2-curvature directed force K. Consider the curve $k(t)$ of initial points on R^2. Then the point $h(t)$ and the angle $\theta(t)$ are constructed exactly as for the Lie algebra $so(3)$ i.e. $h(t)$ is the point of intersection of the two straight lines $\mu(0)$ and $\mu(t)$ tangent to k at the points $k(0)$ and $k(t)$ respectively. The angle between $\mu(0)$ and $\mu(t)$ is equal to $\theta/2$. Thus, the solution of the problem has the form

$$X(t) = X(0)\, e^{\theta(t)h(t)} .$$

It is clear that Magnus' equations become (the radius of the sphere tends to infinity)

$$h'_\mu = L \quad \text{and} \quad h'_\pi = R,$$

$$|h'_\mu| = |L| = \frac{\omega y}{2}\, ctg \frac{\theta}{2} ,$$

$$|h'_\pi| = |R| = \frac{\omega y}{2} ,$$

$$\theta' = \omega,$$

where y is the distance betweeen $h(t)$ and $k(t)$, and the angle between the vector h' and the line is equal to $\theta/2$.

10. Sketch of another proof of Theorem 3 for $e(2)$ without using $su(2)$- or $so(3)$-calculations.

This section is based on observations of S.Anisov.

As noted earlier, it follows by a classical theorem that any element of the group $E(2)$ is a rotation by some angle φ about some point (p,q) of the plane or else parallel transport in the direction of some vector v of the plane, so that compositions of these operations result in a single such operation.

It is well-known that the logarithm of the rotation around a point by the angle || is equal to a matrix depending on the point multiplied by $\varphi (mod\ 2\pi)$, and that the logarithm of the shift by the vector v is the directional derivative in the direction of v multiplied by $|v|$, if the shift is thought of as acting on the space of differentiable functions defined on the plane.

The part $e^r(2)$ of the Lie algebra $e(2)$ corresponding to rotations is identified with R^3 using the coordinates (p,q,φ), where φ is the angle of rotation. We suppose that $K(t)$ belongs to $e^r(2)$.

Both curves $h(t)$ and $k(t)$ are constructed on the plane (p,q) as in the previous sections. The moving coordinate system or $(\omega/2)$*-bicycle wheel system* on the 2-plane is also constructed exactly as in previous sections. The velocity of rotation of this system about the point $k(t)$ is equal $(\omega/2)dt/ds$, where s is arc-length on k.The Greenwich meridian of the system is again denoted by $\mu(t)$.

The system $X' = XK$ and control force K are called *2-curvature directed*, if $r(t) = (\omega/2)ds/dt$, where $r(t)$ is the curvature of k.

It can be proved (as above) that the Greenwich meridian $\mu(t)$ is always tangent to the curve k at the point $k(t)$. Consider $H(t) = \theta(t)h(t)$ which is the logarithm of the product integral $\Pi\ e^{k\theta(t)}$.

Statement 5. If $\theta(t)$ belongs to the interval $(0,2\pi)$, then the point $h(t)$ belongs to the Greenwich meridian $\mu(t)$.

Proof: Consider the orthogonal decomposition:

$$h' = h'_\mu + h'_\pi$$

where h'_μ is the projection on μ and h'_π is the projection on the directio orthogonal to μ. It is sufficient to prove that $|h'_\pi|$ is equal to the velocity of the point $h(t)$ on $\mu(t)$ under the rotation of μ around $k(t)$ with velocity $\omega/2$.

Note that if $e^{H(t)}$ is a rotationwhose angle is in the interval $(0,2\pi)$, then $h(t)$ is its unique fixed point on R^2.

Consider the composition which corresponds to the time moment $t + \Delta t$:

$$e^{\Delta t\omega(t)k(t)} e^{\theta(t)h(t)} = e^{\Delta tK} e^{H} = e^{\theta^* h^*}.$$

The rotation $e^{\theta^* h^*}$ is the resulting rotation about the new point $h^* = h(t + \Delta t)$ (close to the point $h = h(t)$) by angle θ^* which is approximately equal to the angle θ (up to $o(\Delta t)$). We want to find the new rotation point h^*. The geometrical picture is the following: the neighborhood of the point $h(t)$ initially is rotated around the point $h(t)$ by angle θ, and then is shifted by parallel transport along the vector which is orthogonal to the meridian $\mu(t)$ and has norm equal to $\omega y\Delta t$. Here $y = |k(t),h(t)|$ is the length of the meridian arc between the points $k(t)$ and $h(t)$.

The new fixed point $h^* = h(t + \Delta t)$ has the following properties:

1) the angle between μ and velocity vector h' is equal to $\theta/2$,

2) $|h(t),h(t + \Delta t)|\ \sin\ (\Delta/2) = \omega y\Delta t/2$.

Since h^* is the fixed point of the new rotation h^*, and because under this rotation the point h goes into a point Q, the segment (h,h^*) goes into a segment (h^*,Q). Because the angle between (h^*,h) and (h^*,Q) is equal to θ, then the angle between (h^*,h) and (h^*,P) is equal to $\theta/2$. Thus, $|h,h^*| = |h^*,Q|$ and the line (k,h^*) intersects the segment (h,Q) at the point P, which is in the middle of the segment (h,Q). Further, since $|h,Q| = \omega y\Delta t$, then $|h,P| = |\Delta t h'_\perp| = |\Delta t h'|\ \sin\ (\theta/2) = \omega y\Delta t/2$.

As Δt tends to zero, then in the limit we obtain:

$$|h'_\perp| = \omega y/2.$$

The velocity of the point h on the meridian under the rotation of $\mu(t)$ about $k(t)$ with velocity $\omega/2$ is also equal to $\omega y/2$. Thus, these two velocities are equal, and we see that point $h(t)$ does not leave the meridian $\mu(t)$ during the motion. The statement is proved.

To finish the proof of Theorem 3, we need to repeat the final arguments which were used in the first proof.

Theorem 3 is proved.

The following table summarizes the connections between the three Lie algebras.

su(2)	so(3)	e(2)
θ	$\theta/2$	$\theta/2$
ω	$\omega/2$	$\omega/2$
1-curvature-directed force K	2-curvature-directed force K	2-curvature directed force K
ω-bicycle-wheel system	$\frac{\omega}{2}$-bicycle-wheel system	$\frac{\omega}{2}$-bicycle-wheel system
$\lvert k,h\rvert = \sin y$	$\lvert k,h\rvert = \sin y$	$\lvert k,h\rvert = y$
$\theta' = \omega \cos y$	$\theta' = \omega \cos y$	$\theta' = \omega$

11. Integration of the equation $X' = XK$ for the Lie algebras $sl(2,R)$ and $sl(2,R)^+$ in the curvature directed case.

The case of the Lie algebras $sl(2,R)$ and $sl(2,R)^+$ is very similar to the case of the Lie algebras $so(3)$ and $su(2)$. This is to be expected because the Lie algebras $su(2)$ and $sl(2,R)$ are the *compact* and *non-compact* real forms of the same complex simple Lie algebra $sl(2,C)$. The group $SL(2,R)$ is a 2-fold covering over the group $SL(2,R)^+$, since

$$SL(2,R)^+ = SL(2,R)/Z_2$$

Theorem 4. The differential equation $X' = XK$ *for the Lie algebras* $sl(2,R)$ *and* $sl(2,R)^+$ *in the 2-curvature directed case and in the 1-curvature directed case respectively can be solved by the procedure of direct geometric integration exactly as for the Lie algebras* $su(2)$ *and* $so(3)$ *but with one difference: we need to replace great circles of the 2-sphere by geodesic lines on the hyperbolic 2-plane. This means that the point* $h(t)$ *is the intersection point of the two geodesic lines on the hyperbolic 2-plane which are tangent to the curve* $k(t)$ *at its starting and end points. The rotation angle* $\theta(t)$ *(respectively,* $\theta/2$*) is the angle between these lines in the case of* $sl(2,R)$ *(respectively,* $sl(2,R)^+$*).*

In the formulas which are used in the proof it is necessary to replace the functions *sin*, *cos*, and *ctg* by their hyperbolic analogs: *sh*, *ch*, and *cth*.

This is summarized in the following table:

$sl(2,R)$	$sl(2,R)^+$
θ	$\theta/2$
ω	$\omega/2$
1-curvature-directed force K	2-curvature-directed force K
ω-bicycle-wheel system	$\frac{\omega}{2}$-bicycle-wheel system
$\lvert k,h\rvert = sh\ y$	$\lvert k,h\rvert = sh\ y$
$\theta' = \theta\ ch\ y$	$\theta' = \omega\ ch\ y$

Let us recall that the group $SL(2,R)$ is the group of symplectic (homogeneous) transformations of the plane (i.e. preserving the area of the plane); the group $SL(2,R)^+$ is a group of proper isometries of the hyperbolic plane.

It is possible to prove Theorem 4 directly, without reference to the general theory of compact and non-compact real forms of simple complex Lie algebras. We sketch a direct proof.

Consider the Lie algebra $sl(2,R)$ which consists of all matrices of the form:

$$X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix},$$

where a,b,c are arbitrary real numbers. Then it is easy to see that $X^2 = m^2 E$, where E is the identity matrix, and $m^2 = a^2 + bc$. Thus, if the matrix H belongs to $sl(2,R)$, then we can always represent it in the form:

$$H = \theta h,$$

where $\theta^2 = -\det H = a^2 + bc$ and $\det h = -1$. It follows from the formula $H^2 = \theta^2 h$ that

$$e^H = E\ ch\ \theta + h\ sh\ \theta.$$

This formula is similar to the formula for the case $su(2)$. Let us identify Lie algebra $sl(2,R)$ with 3-space R^3 with coordinates a,b,c. Then the pseudo-sphere is determined by the equation:

$$\theta^2 = -1, \text{ i.e. } a^2 + bc = -1.$$

The equation $a^2 + bc = -1$ determines the 2-dimensional hyperboloid in 3-space. After the transformation: $b = b' + c'$ and

$c = b' - c'$, we obtain the equation $a^2 + b'^2 - c'^2 = -1$. Thus, the curves k and h are the curves on the hyperboloid which represents the hyperbolic plane (realized in a standard way in pseudo-Euclidean 3-space).

To get the proof of Theorem 4, we repeat the argument for *su(2)* with the changes mentioned earlier.

REFERENCES

1 Chacon, R.V., and Fomenko, A.T., Recursion formulas for the Lie integral, Advances in Math. v.88, No. 2, 1991, pp.200-257.

2 Chacon, R.V., and Fomenko, A.T., Stokes' formula for Lie algebra valued connection and curvature forms, Advances in Math. vol.88, No.2, 1991, pp. 258 - 300.

3 Dollard, J.D., and Friedman, C.N., *Product Integration*, Addison-Wesley, 1979.

4 Feynman R.C., An operator calculus having applications in quantum electrodynamics, Phys. Rev. 84 (2), 1951, pp.108-128.

5 Bialyncki-Birula, I., Mielnik, B., and Plebanski, J., Explicit solution of the continuous Baker-Campbell-Hausdorff problem and a new expression for the phase operator, Ann.Phys. 51 (1969), pp. 187-200.

6 Karasev, M. V., and Mosolova, M.V., The infinite products and T-products of exponents, J.Theoret.Math.Phys. 28 (2) (1976),
7 pp.189-200.

7 Chen, K.T., Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula, Ann.of Math. 65 (1957), pp.163-178.

8 Fomenko, A.T., and Chacon, R.V., Recurence formulas for the homogeneous terms of the converging expansion representing the logarithm of the product integral on Lie groups, Functional Anal. and its Applications, 1990, v.24, No. 1, pp.48 - 58. (See corresponding English translation).

9 Chacon, R.V. and Fomenko, A.T. Integration of the connection equation for arbitrary Lie algebras, to appear.

Department of Mathematics
University of British Columbia
Vancouver, BC, Canada V6T 1Y4

Received July 28, 1994

Department of Topology and Geometry
Moscow State University
119899 Moscow, Russia

C.R. Math. Rep. Acad. Sci. Canada - Vol. XVI, No. 4, August 1994 août

COVERING THEOREMS FOR OPEN CONTINUOUS MAPPINGS HAVING TWO VALENCES

ABDALLAH LYZZAIK
Presented by J.S. Halperin, F.R.S.C.

Abstract

Let $\mathcal{X}$ be an open Riemann surface with finite genus and finite number of boundary components, and let $\mathcal{Y}$ be a closed Riemann surface. An open continuous function from $\mathcal{X}$ to $\mathcal{Y}$ is termed a (p,q)−map, $0 < q < p$, if it has a finite number of branch points and assumes every point in $\mathcal{Y}$ either p or q times, counting multiplicity, with possibly a finite number of exceptions. These comprise the most general class of all nontrivial functions having two valences between $\mathcal{X}$ and $\mathcal{Y}$. The object of this paper is to announce the most recent results involving the geometric and combinatorial relations governing (p,q)−maps.

INTRODUCTION

The covering properties of open continuous functions having essentially two valences p and q $(0 < q < p)$, termed (p,q)−maps, have generated considerable interest in the past two decades (see [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].) Recently, Lyzzaik and Stephenson [4] studied (p,q)−maps from the open unit disc to closed (compact) Riemann surfaces and concluded in a unified way the earlier results, among others, on (p,q)−maps of the unit disc. More general settings, involving (p,q)− maps from open Riemann surfaces of finite genus and finite number of boundary components into closed Riemann surfaces, were treated by Srebro and Wajnryb [9, 10] but only in the case where $q = p - 1$. Restrictions imposed on these maps vis-à-vis their domains as in [2, 4, 8] or their valences as in [9, 10] were deliberate and somehow prompted by the limitations of the already established techniques.

In this paper we consider (p,q)−maps from open Riemann surfaces of finite genus and finite number of boundary components into closed Riemann surfaces, allowing branch points and exceptional points. A point of the image surface of a (p,q)−map will be called *exceptional* if it is not assumed p or q times, counting multiplicity. Accordingly, our (p,q)−maps will be far more general than any that have been studied so far. Indeed, these are about the

[0]1991 *Mathematics Subject Classification* (1985 *Revision*). Primary 57M10, 14E25, 30F10.

most general that one may consider in view of the desired and expected results. Let us first formulate the precise definition of (p,q)-maps.

Definition. *Let p and q be integers satisfying $0 < q < p$, $\mathcal{X}$ an open Riemann surface of finite genus and finite number of boundary components, and $\mathcal{Y}$ a closed Riemann surface. A function $f : \mathcal{X} \to \mathcal{Y}$ is termed a (p,q)-map if f satisfies the following conditions:*

(a) f is open and continuous.

(b) f admits every point of $\mathcal{Y}$, with at most finitely many exceptions, either p or q times, counting multiplicity, with at least one point admitted exactly p times.

(c) f admits only a finite number of branch points.

STATEMENTS OF RESULTS

It is apparent from the literature that the success of the study of the covering properties of a (p,q)-map f depends mostly on whether one can establish an embedding theorem whereby f lifts to an embedding into a p-fold covering of the image surface. This idea was first introduced by Lyzzaik and Styer [5] to settle conjectures involving analytic $(p,p-1)$-maps. Soon, the idea was adopted by Srebro [8] and by Srebro and Wajnryb [9] in treating meromorphic $(p,p-1)$-maps and $(p,p-1)$-maps between Riemann surfaces, respectively. Subsequently, the idea was extended independently by Brannan and Lyzzaik [2] and Carne, Ortel and Smith [3] to locally-univalent (p,q)-maps between the unit disc and the Riemann sphere. With the same desire, Lyzzaik and Stephenson [4] considered the broader class of (p,q)-maps from the unit disc to closed Riemann surfaces, allowing finite sets of branch points and exceptional points, thus establishing all the earlier results on (p,q)-maps of the unit disc. In the latter paper the authors observed that their so-called embedding theorem may extend to more general (p,q)-maps (see example 5 of [4]), and they asked whether the theorem can generalize to (p,q)-maps of multiply-connected domains. Our result in this direction states the following:

Theorem. *Let $f : \mathcal{X} \to \mathcal{Y}$ be a (p,q)-map. Then there exists a p-fold covering surface $(\tilde{\mathcal{Y}},\pi)$ of $\mathcal{Y}$ and an embedding $\phi : \mathcal{X} \to \tilde{\mathcal{Y}}$ so that $f \equiv \pi \circ \phi$.*

The theorem allows us to view the image surface of a (p,q)-map as a subset of a p-fold covering of the the image surface. Working in this setting, we invoke the special type of homotopy, termed *(p,q)-homotopy*, developed in [4] to perform a continuous deformation in $\tilde{\mathcal{Y}}$ of the subsurface $\phi(\mathcal{X})$ that turns $\phi(\mathcal{X})$ into a polygonal subsurface of $\tilde{\mathcal{Y}}$ having an empty exterior and the same "valence structure" as that of $\phi(\mathcal{X})$. This surface will work as the image surface of a $(p,q)-$map $h : \mathcal{X} \to \mathcal{Y}$ that is " $(p,q)-$homotopic"to f. At this point it becomes apparent that it would be useful for some boundary components of $\mathcal{X}$ to assume points for their ideal boundaries verses contours to others. This yields a restricted compactification, $\mathcal{X}^{*}$, of $\mathcal{X}$, termed *admissible compactification* of $\mathcal{X}$, which depends solely on f. Using $\mathcal{X}^{*}$, we show that h extends continuously to a simplicial map from a triangulation of $\mathcal{X}^{*}$ to a triangulation of $\mathcal{Y}$.

Theorem. *Let $f : \mathcal{X} \to \mathcal{Y}$ be a $(p,q)-$map. Then f is $(p,q)-$homotopic to a $(p,q)-$map $h : \mathcal{X} \to \mathcal{Y}$ that extends continuously to a simplicial map from a triangulation of an admissible compactification of $\mathcal{X}$ to a triangulation of $\mathcal{Y}$.*

By virtue of this theorem it will be sufficient to consider simplicial $(p,q)-$maps for the study of the valence structure of general $(p,q)-$maps. For this purpose, we require some notation and terminology. Because a $(p,q)-$map $f : \mathcal{X} \to \mathcal{Y}$ is open and continuous, a *branch point* and the *order of a branch point* will mean the same for f as for an analytic function (see [11].) The *total branch order* of f will be denoted by β_f. A companion of f is the integer-valued function $\nu_f(y)$ defined as the number of times, counting multiplicity, that f assumes a value y in $\mathcal{Y}$. Set

$$\mathcal{P} = \{y \in \mathcal{Y} : \nu_f(y) = p\},$$

$$\mathcal{Q} = \{y \in \mathcal{Y} : \nu_f(y) = q\}$$

and

$$\mathcal{E} = \{y \in \mathcal{Y} : \nu_f(y) \neq p, q\}.$$

We call $\mathcal{P}$ the *p-set*, $\mathcal{Q}$ the *q-set*, and $\mathcal{E}$ the *exceptional set*, of f. Evidently, these sets partition $\mathcal{Y}$. The $p-$set happens to be open and to have finitely many connected components each with finitely many boundary components. Every point $e \in \mathcal{E}$, called *exceptional point*, will be assigned the value

$$\delta_f(e) = q - \nu_f(e)$$

which we call the *deficiency* of f at e. Generally speaking, $\delta_f(e)$, may, unlike in [4], take negative values. This fortunately can be avoided by normalizing our $(p,q)-$maps without loosing generality; the normalized maps will be termed *normal* $(p,q)-$*maps*. For these maps the *total deficiency* of f is defined as the sum of all the deficiences $\delta_f(e)$. We will also utilize some basic concepts and results from surface topology. For instance, the *genus*, *number of boundary components* and *Euler's characteristic* of a surface, denoted by g(.), $\eta(.)$ and $\chi(.)$, respectively, will be used in connection with the Riemann-Hurwitz relation to establish the following theorem:

Theorem. *Let* $f : \mathcal{X} \to \mathcal{Y}$ *be a normal* $(p,q)-$*map. Then*

$$\beta_f + \delta_f = 2[(g(\mathcal{X}) - 1) - q(g(\mathcal{Y}) - 1)] + \eta(\mathcal{X}) + (p-q)\chi(\mathcal{P}).$$

This theorem is the key tool for all the combinatorial results of the paper. For instance, we have:

Theorem. *Let* $f : \mathcal{X} \to \mathcal{Y}$ *be a normal* $(p,q)-$*map. Then*

$$\beta_f + \delta_f \geq 2[(g(\mathcal{X}) - 1) - p(g(\mathcal{Y}) - 1)] - (p-q-1)\eta(\mathcal{X}).$$

Further, this inequality is best possible.

If $\mathcal{P}$ is simply-connected, then we conclude:

Theorem. *Let* $f : \mathcal{X} \to \mathcal{Y}$ *be a normal* $(p,q)-$*map. If* $\mathcal{P}$ *is simply- connected, then*

$$\beta_f + \delta_f \geq 2[(g(\mathcal{X}) - 1) - q(g(\mathcal{Y}) - 1)] + p - q + \eta(\mathcal{X}).$$

Further, this inequality is best possibe.

If $\mathcal{X}$ is the unit disc, **D**, the theorem gives Theorem 5 of [4].

We pay special attention to normal $(p,q)-$maps having zero deficiency, termed *normal* $BQ-$*maps*, in consistence with [4].

Theorem. *Suppose* $f : \mathcal{X} \to \mathcal{Y}$ *is a normal* $BQ-$*map. Then*

$$a \leq \beta_f \leq b,$$

where

$$a = 2[(g(\mathcal{X}) - 1) - p(g(\mathcal{Y}) - 1)] - (p-q-1)\eta(\mathcal{X})$$

and

$$\begin{aligned} b &= \min\{2[(g(\mathcal{X})-1)-q(g(\mathcal{Y})-1)]+(p-q+1)\eta(\mathcal{X}), \\ &\qquad (p-q+1)(2[(g(\mathcal{X})-1)-q(g(\mathcal{Y})-1)]+(p-q)\eta(\mathcal{X}))\}. \end{aligned}$$

The inequalities are best possible when $q = p-1$ *or*

$$b = 2[(g(\mathcal{X})-1)-q(g(\mathcal{Y})-1)]+(p-q+1)\eta(\mathcal{X}).$$

In case $q = p-1$, this theorem is essentially a result of Srebro andWajnryb (see [9, Covering Theorem 1]). Using this result we conclude:

Theorem. *A necessary condition for a normal BQ−map* $f : \mathcal{X} \to \mathcal{Y}$ *to exist is that*

$$g(\mathcal{X}) \geq q(g(\mathcal{Y})-1)+1-(p-q)\eta(\mathcal{X})/2.$$

Working with $(p, p-1)$−maps, Srebro and Wajnryb [9] proved that the condition is also sufficient; else their result is a special case of the above theorem.

In the special case when $\eta(\mathcal{X}) = 1$, an interesting geometric characterization of normal BQ−maps yields:

Theorem. *Suppose* $f : \mathcal{X} \to \mathcal{Y}$ *is a normal BQ−map, and* $\eta(\mathcal{X}) = 1$. *Then* $g(\mathcal{X}) \geq g(\mathcal{Y})$ *and*

$$\beta_f = 2[(g(\mathcal{X})-1)-p(g(\mathcal{Y})-1)]-p+q+1+2n(p-q)$$

where n is an integer satifying, $0 \leq n \leq g(\mathcal{Y})$.

To illustrate its use, if $\mathcal{X} = \mathbf{D}$ then n and $g(\mathcal{Y})$ must be zero and $\beta_f = p+q-1$; a result which can be found in [4].

References

[1] D. A. Brannan and W. E. Kirwan, *Some covering theorems for analytic functions,* J. London Math. Soc. (2)**19** (1979), 93-101.

[2] D. A. Brannan and A. K. Lyzzaik, *Some covering properties of locally-univalent functions,* Ann. Acad. Sci. Fenn. Ser. A I Math. **13** (1988), 3-23.

[3] K. Carne, M. Ortel and W. Smith, *Multiplicities of locally one-to-one continuous maps of the plane,* Bull. London Math. Soc. **18** (1987), 438-442.

[4] A. K. Lyzzaik and K. Stephenson, *The structure of open continuous mappings having two valences,* Trans. Amer. Math. Soc. Vol. **327**, No. **2** (1991), 525-566.

[5] A. K. Lyzzaik and D. Styer, *A covering surface conjecture of Brannan and Kirwan,* Bull. London Math. Soc. **14** (1982), 39-42.

[6] M. Ortel and W. Smith, *A covering theorem for continuous locally univalent maps of the plane,* Bull. London Math. Soc. **18** (1986), 359-363.

[7] U. Srebro, *Deficiences of immersions,* Pacific J. Math. **113** (1984), 493-496.

[8] U. Srebro, *Covering theorems for meromorphic functions,* J. Analyse Math. **44** (1984/85), 235-250.

[9] U. Srebro and B. Wajnryb, *Covering theorems for Riemann surfaces,* J. Analyse Math. **46** (1986), 283-303.

[10] U. Srebro, *Covering theorems for open surfaces,* Proceedings of the 1985 Georgia Topology Conference: Geometry and Topology, Dekker, Amsterdam, 1987, pp. 265-275.

[11] S. Stoïlow, *Principes topologiques de la theorie des fonctions analytiques,* Gauthier-Villars, Paris, 1938.

Department of Mathematics, American University of Beirut, Beirut,Lebanon.

Received July 5, 1994

Comparing two topological invariants of homogeneous complex manifolds

Bruce Gilligan

Presented by J. Arthur, F.R.S.C.

Abstract. We compare two topological invariants of a manifold X, the number of its ends $e(X)$ and the codimension d_X of its top nonvanishing homology group with coefficients in $\mathbb{Z}_2$. In general, $e(X) > 1 \Longrightarrow d_X = 1$. The Moebius band provides a counterexample to the converse. But if $X = G/H$ is a homogeneous complex manifold and X is Kähler, H has a finite number of connected components, or $\mathcal{O}(X) \neq \mathbb{C}$, then we show that $d_X = 1 \Longleftrightarrow e(X) = 2$.

1 Introduction

For any manifold X the notion of ends (in the sense of Freudenthal [7]) is defined. Let $e(X)$ denote the number of ends of X. Another topological invariant of X can be defined in the following way. If X has dimension n, set

$$d_X := \min\{ r \mid H_{n-r}(X, \mathbb{Z}_2) \neq 0 \}.$$

By Poincaré duality d_X is dual to the non-compact dimension $n_c(X; \mathbb{Z}_2)$ of X with respect to $\mathbb{Z}_2$ introduced by H. Abels in [1]. In [1] it is proved that if X is a connected, locally compact, paracompact space with $e(X) > 1$, then $d_X = 1$. However, the Moebius band M satisfies $d_M = 1$ and $e(M) = 1$, thus yielding a counterexample to the converse.

The purpose of this note is to point out that $d_X = 1$ does imply $e(X) > 1$ in certain settings involving homogeneous complex manifolds. For our present purposes a complex manifold X is *homogeneous* if there is a Lie group G which is acting transitively on X as a group of biholomorphic transformations. In this situation it is well-known that $X = G/H$, where H is a closed subgroup of G and the coset space G/H has an invariant complex structure. What we will prove is the following.

Theorem *Suppose $X = G/H$ is a homogeneous complex manifold. Assume that one of the following conditions holds:*

1) X is Kähler.

2) H is a closed subgroup of G having a finite number of connected components.

3) $\mathcal{O}(X) \neq \mathbb{C}$.

Then $e(X) = 2$ if and only if $d_X = 1$.

There is a description in each of these three settings of those homogeneous complex manifolds which have two ends: see [12], [11], [8], and [13]. In [10] it is proved that if $X = G/H$ satisfies $e(X) > 1$, then there is a complex Lie group acting transitively on X.

2 Some Preliminaries

Using a spectral sequence argument, we noted the following in [2].

Lemma 1 *1) Suppose $X \xrightarrow{F} B$ is an orientable fiber bundle (e.g., if $\pi_1(B) = 0$), then $d_X = d_F + d_B$.*

2) If $X \xrightarrow{F} B$ is a fiber bundle and B has the homotopy type of a CW-complex of dimension q, then $d_X \geq d_F + (\dim B - q)$. Further, if B is homotopy equivalent to a compact manifold, then $d_X \geq d_F + d_B$.

3) Suppose X and Y are connected manifolds, $\pi : X \to Y$ is an unramified covering, $M \subset Y$ is a compact submanifold and Y is retractable onto M. Then $d_X = d_Y + d_{\pi^{-1}(M)}$. In particular, if π is infinite, then $d_X > d_Y$.

Proofs will proceed by induction on $n := \dim_{\mathbb{C}} X$. If $n = 1$ and X is homogeneous, then

$$d_X = 1 \Longrightarrow X = \mathbb{C}^* \Longrightarrow e(X) = 2. \tag{1}$$

If X is a Kähler or holomorphically separable manifold and F is a complex submanifold of X, then F satisfies the same property. As well, if $G/H \to G/J$ is a homogeneous fibration, then condition 2) of the theorem is inherited by the fiber J/H. Thus the following observations simplify the induction step.

Lemma 2 *Suppose $X \xrightarrow{F} B$ is a fiber bundle with connected positive dimensional fiber and base and $d_X = 1$. Assume induction applies, i.e., $d_F = 1 \Longrightarrow e(F) > 1$. Then in order to prove $e(X) \geq 2$, we may assume F is compact.*

PROOF: If F is not compact, then

$$d_X \geq d_F + (\dim B - q)$$

implies $d_F = 1$ and applying induction to the fiber one has $e(F) > 1$. Also from above we see that $q = \dim B$, where B has the homotopy type of a CW-complex of dimension q. It follows from this that B must be compact, since Munkres [15] proved that for B a connected, non-compact triangulated manifold there is a subcomplex of B of dimension $q < \dim B$ that is a deformation retract of B. But this would contradict the above formula. Thus B is compact and hence $e(X) \geq e(F) > 1$ follows from the proof of Lemma 2 in [8]. □

Lemma 3 *Suppose $\pi : X \xrightarrow{F} B$ is a locally trivial fiber bundle with connected compact fiber F and X (equivalently B) is connected and noncompact. Then the induced map $\pi_c^* : H_c^1(B) \to H_c^1(X)$ is an isomorphism. In particular, $d_X = 1 \Longrightarrow d_B = 1$.*

PROOF: The E_2-term of the spectral sequence for cohomology with compact supports has two terms of total degree one. These are $H^1_c(B; \mathfrak{H}^0_c(F)) = H^1_c(B)$ and $H^0_c(B; \mathfrak{H}^1_c(F))$. Because B is connected and noncompact and the sheaf $\mathfrak{H}^1_c(F)$ is locally constant, the zero section is the only section of $\mathfrak{H}^1_c(F)$ with compact support. Thus $H^0_c(B; \mathfrak{H}^1_c(F)) = 0$. Since all the differentials d_r to and from terms of E_r with total degree one are zero, we see that $H^1_c(X) \cong H^1_c(B)$ and the natural map π_c induces the isomorphism. □

For any homogeneous complex manifold G/H there exists a closed subgroup I of G containing H such that the induced fibration $\pi : G/H \to G/I$ is the *holomorphic reduction* of G/H, i.e., G/I is holomorphically separable and $\pi^*(\mathcal{O}(G/I)) \cong \mathcal{O}(G/H)$. If the fibers of the holomorphic reduction of G/H are discrete, then G/H is said to satisfy the *maximal rank condition.*

2.1 The Kähler case

Proposition 1 *Suppose G/H is a Kähler homogeneous manifold with $d_{G/H} = 1$. Then $e(G/H) = 2$.*

PROOF: The proof will proceed by induction on $n := \dim_{\mathbb{C}} X$. For $n = 1$, see (1). If G is a complex Lie group (resp. real Lie group), let $G/H \to G/J$ be the normalizer fibration of G/H (resp. the $\mathfrak{g}$-anticanonical fibration). Assume first that $\dim G/H > \dim J/H \geq 0$. Let $\tilde{J}$ consist of those connected components of J which meet H. By Lemma 2 we may assume that $\tilde{J}/H$ is compact. The result then follows from part B) below. Otherwise, $\tilde{J} = G$. Then G/H can be written as the quotient of a complex Lie group by a discrete subgroup [13, p. 64] and the result is given by part A) below.

We still have to prove the following: Suppose X is a homogeneous complex manifold with $d_X = 1$ and one of the following conditions holds for X:

A) $X = G/\Gamma$ is Kähler, where G is a complex Lie group and Γ is a discrete subgroup of G.

B) $X = G/J$ is an orbit in some $\mathbb{P}^m$.

Then $e(X) = 2$.

A): Let $G = R \cdot S$ be a Levi decomposition of G with $\dim R > 0$. If the R-orbits in G/Γ are not closed, then there exists a closed complex subgroup N of G with $\Gamma R \subset N$ and $N \neq G$, e.g., see [12, Prop. 2, p. 168]. Letting $\tilde{N} := N^\circ \cdot \Gamma$, by Lemma 2 we may assume that the connected fiber $\tilde{N}/\Gamma$ is compact. Then $d_{G/\tilde{N}} = 1$ and because $\dim R > 0$, the result follows by induction and the fact that the Kähler form "pushes down" to $G/\tilde{N}$.

Now assume that the R-orbits are closed in G/Γ. By Lemma 2 we may assume $R\Gamma/\Gamma$ is compact. Then $G/R\Gamma = S/\Lambda$, where S is a maximal semisimple complex Lie group in G and $\Lambda := R\Gamma \cap S$. Since the Kähler form of G/Γ "pushes down" to S/Λ it follows that Λ is algebraic [4] and hence finite. From Lemma 1 one has $1 = d_{G/H} = d_{S/\Lambda} = d_S = \dim_{\mathbb{C}} S \geq 3$, a contradiction. This argument also handles the case $G = S$. For G solvable, we proceed as in [12]. Let $G/\Gamma \to G/I$ be the holomorphic reduction, if $\mathcal{O}(G/\Gamma) \neq \mathbb{C}$.

Then G/I is Stein with $d_{G/I} = 1$ and thus $G/I = \mathbb{C}^*$. By Lemma 1 we may assume that the fiber I/Γ is compact and so $e(G/\Gamma) = 2$. If $\mathcal{O}(G/\Gamma) = \mathbb{C}$, then by [16] we have that G is abelian and G/Γ is a Cousin group. Since $d_{G/\Gamma} = 1$, it is clear that $e(G/\Gamma) = 2$.

B): Let $G/J \to G^{\mathbb{C}}/\hat{J}$ be the imbedding of G/J into the orbit of the complexification $G^{\mathbb{C}}$ of G. Since $(G^{\mathbb{C}})'$ acts as an algebraic group on projective space, by Chevalley's theorem [6], it follows that the $(G^{\mathbb{C}})'$–orbits are closed and thus we have a fibration $G^{\mathbb{C}}/\hat{J} \to G^{\mathbb{C}}/\hat{J}\cdot(G^{\mathbb{C}})'$. If $\dim G^{\mathbb{C}}/\hat{J}\cdot(G^{\mathbb{C}})' > 0$, then $G^{\mathbb{C}}/\hat{J}\cdot(G^{\mathbb{C}})' = \mathbb{C}^*$. Since the induced fibration of X also fibers onto $\mathbb{C}^*$ with compact connected fiber, it is clear $e(X) = 2$. Next assume $(G^{\mathbb{C}})'$ is transitive on $G^{\mathbb{C}}/\hat{J}$. Consider the fibration of $G^{\mathbb{C}}/\hat{J}$ induced by the radical $R_{(G^{\mathbb{C}})'}$ of $(G^{\mathbb{C}})'$. By Lemma 2 we may assume that the fiber of the induced fibration of G/J is compact. But the orbits of $R_{(G^{\mathbb{C}})'}$ in $\mathbb{P}^m$ are holomorphically separable, by Lie's theorem. This would contradict the maximum principle, unless $R^{\mathbb{C}} \subset \hat{J}$. Because we may assume that the action of G on G/J is almost effective, it follows that G is semisimple.

Now $\mathfrak{g}\cap i\mathfrak{g}$ is an ideal in $\mathfrak{g}$ and the corresponding connected subgroup M of G consists of the product of some of the simple factors of G. Consider the fibration $G/J \to G/J\cdot M$. By Lemma 2 we may assume that $J\cdot M/J$ is compact. By Lemma 3 one has $d_{G/J\cdot M} = 1$. Since $G/(J\cdot M)^\circ \to G/(J\cdot M)$ is a finite covering, it follows from Lemma 1 that $d_{G/(J\cdot M)^\circ} = 1$. Thus $e(G/(J\cdot M)^\circ) = 2$. From [13, Prop. 8, p. 75] we see that $e(G/J\cdot M) = 2$ and so G/H also has two ends. Thus we may assume that G is a real form of $G^{\mathbb{C}}$ and G/J is the quotient of (real) algebraic groups. Then $d_{G/J} = 1 \Longrightarrow d_{G/J^\circ} = 1 \Longrightarrow e(G/J^\circ) = 2$. Again the result follows from Proposition 8 in [13]. □

2.2 The case $|H/H^\circ| < \infty$

Proposition 2 *Suppose $X = G/H$ is a homogeneous complex manifold with $d_X = 1$ and $|H/H^\circ| < \infty$. Then $e(X) = 2$.*

PROOF: Since $G/H^\circ \to G/H$ is a finite covering, it follows that $d_{G/H^\circ} = 1$. Now G/H° is homeomorphic to $K/L \times \mathbb{R}$, where K (resp. L) is a maximal compact subgroup of G (resp. of H, contained in K) by Borel's theorem [5] and thus $e(G/H^\circ) = 2$. In the case when G is a complex Lie group it follows, again because H has only finitely many connected components, that $e(G/H) = 2$, see Proposition 1 in [9].

If G is a real Lie group, let $G/H \to G/J$ be the $\mathfrak{g}$–anticanonical fibration of G/H. If $\dim G/J > 0$, then by Lemma 2, we may assume that $\tilde{J}/H$ is compact, where $\tilde{J} := J^\circ\cdot H$. The result then follows by induction from part B) of Proposition 1, if we note that $d_{G/J} = 1$ by Lemma 3. Otherwise, $J = G$, and so G/H can be written as the quotient of the complex Lie group G/H° by the finite subgroup H/H°. This is a special case of the above proof for complex groups. □

2.3 The case $\mathcal{O}(G/H) \neq \mathbb{C}$

The proof of the equivalence of $d_{G/H} = 1$ and $e(G/H) = 2$ if $\mathcal{O}(G/H) \neq \mathbb{C}$ and G and H are complex Lie groups was given in [2]. For G and H real Lie groups with $\mathcal{O}(G/H) \neq \mathbb{C}$

the proof could be based on a modification of Proposition 3 in [2]. For the reader's convenience, we present a direct proof when $d_{G/H} = 1$. This argument also works for complex groups if one substitutes the normalizer fibration for the $\mathfrak{g}$-anticanonical in the following. In order to set this up we introduce some notation. Suppose G is a (connected real) Lie group and H is a closed subgroup such that $X := G/H$ has a G-invariant complex structure. Assume $\mathcal{O}(X) \neq \mathbb{C}$. Let $G/H \to G/I$ be the holomorphic reduction of G/H and set $\tilde{I} := H \cdot I^{\circ}$. Then the fibration $G/H \to G/\tilde{I}$ has connected fiber $\tilde{I}/H$ and its base $G/\tilde{I}$ satisfies the maximal rank condition. Let $G/I \to G/J$ be the $\mathfrak{g}$-anticanonical fibration of G/I and set $\tilde{J} := \tilde{I} \cdot J^{\circ}$. Recall that $L := J^{\circ}/I^{\circ}$ is a complex Lie group [13] and so $\tilde{J}/\tilde{I}$ can be written in the form L/Γ, where $\Gamma := \tilde{I}/I^{\circ}$ is a discrete subgroup of L.

Proposition 3 *Suppose G is a connected Lie group and H is a closed subgroup such that G/H has a (left) G-invariant complex structure. Assume $d_{G/H} = 1$ and $\mathcal{O}(G/H) \neq \mathbb{C}$. Then, with the above notation, either $\tilde{J} = \tilde{I}$ and the $\mathfrak{g}$-anticanonical fibration of G/I is a covering or else $L/\Gamma = \mathbb{C}^*$ and $G/\tilde{J}$ is a homogeneous projective rational manifold. In either case, $e(G/H) = 2$.*

PROOF: Note that if $\tilde{J}/H$ is compact, then so is $\tilde{J}/\tilde{I}$. Since $G/\tilde{I}$ satisfies the maximal rank condition, it follows that $\tilde{J} = \tilde{I}$. The result is now a consequence of Proposition 1 applied to the fibration $G/H \to G/J$ which has compact connected fiber $\tilde{I}/H$.

We assume throughout the rest of the proof that $\tilde{J}/\tilde{I}$ is not compact. Thus $\dim L > 0$. If L is solvable, then L/Γ is homotopy equivalent to a compact submanifold. It follows from the fibration lemma that $d_{L/\Gamma} = d_{G/H} = 1$. Since L/Γ is a holomorphically separable complex solv-manifold, L/Γ is Stein [14]. It is well-known that if X is a connected Stein manifold, then $\dim_{\mathbb{C}} X \leq d_X$. Hence $d_{L/\Gamma}$ is one-dimensional and so is biholomorphic to $\mathbb{C}^*$. But then $e(G/H) = e(G/\tilde{I}) \geq e(L/\Gamma) = 2$, where the inequality follows from Lemma 2 in [8] and the fact that $G/\tilde{I} \overset{L/\Gamma}{\to} G/\tilde{J}$ has compact base. The result now follows from Corollary 11, p. 80 in [13].

We claim that L cannot be mixed or semisimple. Assume first that L is mixed and, without loss of generality, simply connected. Let $L = S \cdot R$ be a Levi decomposition of L with $\dim R > 0$. Then by Theorem 2 in [8] there exists a closed proper complex subgroup M of L containing R and Γ. By Lemma 2 we may assume M/Γ is compact. Since $G/\tilde{I}$ satisfies the maximal rank condition, $G/\tilde{I}$ cannot contain any positive dimensional compact complex submanifolds. This is a contradiction. Thus $R = \{e\}$ and $L = S$ is semisimple. We now note why this is also impossible. For, then Γ would be algebraic [3] and thus finite. In particular, S/Γ would be homotopy equivalent to a compact submanifold. From Lemma 1 one would have $1 = d_{G/H} \geq d_{S/\Gamma} = d_S = \dim_{\mathbb{C}} S \geq 3$, a contradiction. □

Acknowledgements: This work was partially supported by NSERC Grant A3494. We also thank H. Abels for helpful discussions concerning this work.

References

[1] H. Abels, Some topological aspects of proper group actions; non-compact dimension of groups, *J. London Math. Soc.* (2), 25 (1982), 525–538.

[2] D. N. Akhiezer and B. Gilligan, On complex homogeneous spaces with top homology in codimension two, *Canad. J. Math.*, to appear.

[3] W. Barth and M. Otte, Invariante holomorphe Funktionen auf reduktiven Liegruppen, *Math. Ann.* 201 (1973), 97–112.

[4] F. Berteloot and K. Oeljeklaus, Invariant plurisubharmonic functions and hypersurfaces on semi-simple complex Lie groups, *Math. Ann.* 281 (1988), 513–530.

[5] A. Borel, Les bouts des espaces homogènes de groupes de Lie, *Ann. of Math.* 58 (1953), 443–457.

[6] C. Chevalley, *Théorie des groupes de Lie II : Groupes algébriques.* Hermann, Paris, 1951.

[7] H. Freudenthal, Über die Enden topologischer Räume und Gruppen, *Math. Z.* 33 (1931), 692–713.

[8] B. Gilligan, Ends of homogeneous complex manifolds having nonconstant holomorphic functions, *Arch. Math.* 37 (1981), 544–555.

[9] B. Gilligan, On the ends of complex manifolds homogeneous under a Lie group, *Proc. Sympos. Pure Math., 52, Part 2* (1991), Part 2, 217–224.

[10] B. Gilligan and P. Heinzner, Lifting holomorphic actions to bundles, to appear.

[11] B. Gilligan and A.T. Huckleberry, Complex homogeneous manifolds with two ends, *Michigan Math. J.* 28 (1981), 183–198.

[12] B. Gilligan, K. Oeljeklaus and W. Richthofer, Homogeneous complex manifolds with more than one end, *Canad. J. Math.* 41 (1989), 163–177.

[13] A. T. Huckleberry and E. Oeljeklaus, Classification Theorems for Almost Homogeneous Spaces, *Publ. Inst. E. Cartan* 9, Nancy, 1984.

[14] A. T. Huckleberry and E. Oeljeklaus, On holomorphically separable complex solvmanifolds, *Ann. Inst. Fourier* 36 (1986), 57–65.

[15] James R. Munkres, A theorem on open manifolds, preprint.

[16] K. Oeljeklaus and W. Richthofer, On the structure of complex solvmanifolds, *J. Differential Geom.* 27 (1988), 399–421.

Dept. of Math. and Stats., Univ. of Regina
Regina, Canada S4S 0A2

e-mail: gilligan@max.cc.uregina.ca

Received July 7, 1994

C.R. Math. Rep. Acad. Sci. Canada - Vol. XVI, No. 4, August 1994 aout

CONTINUOUS WAVELET TRANSFORMS ON SEMISIMPLE LIE GROUPS AND ON CARTAN MOTION GROUPS

K. TRIMECHE

Presented by G.A. Elliott. F.R.S.C.

Abstract.
In this work we define and study wavelets and continous wavelet transforms on semisimple Lie groups G of real rank ℓ and on $p = G_0/K$ where G_0 are Cartan motion groups associated with $X = G/K$. Next we prove for these transforms Plancherel and inversion formulas, and using Abel transforms on G and p, we give relationships between these wavelets on G and p (resp. these continuous wavelet transforms on G and p) and classical wavelets on $\mathbb{R}^\ell$ (resp. the classical continuous wavelet transform on $\mathbb{R}^\ell$).

I. Wavelets and continuous wavelet transform on G.

Let G be a noncompact connected real semisimple Lie group with finite center, $\mathfrak{g}$ the Lie algebra of G. Let $\mathfrak{g} = k + p$ be the Cartan decompsition of $\mathfrak{g}$. Let $\mathfrak{a} \subset p$ be a maximal abelian subspace, $\mathfrak{a}^*$ its (real) dual. The dimension ℓ of $\mathfrak{a}$ is called the real rank of G. We can identify $\mathfrak{a}$ with $\mathbb{R}^\ell$. We denote by Σ the set of restricted roots. Let W be the Weyl group associated with Σ and $|W|$ its cardinality. Let Σ^+ be the set of positive roots, and put $\Sigma_0^+ = \Sigma^+ \cap \{\alpha \in \Sigma \,/\, \frac{1}{2}\alpha \notin \Sigma\}$. Put $\rho = \frac{1}{2}\sum_{\alpha \in \Sigma^+} m_\alpha . \alpha$, with m_α the multiplicity of α. Let K (resp. A) be the analytic subgroup of G with Lie algebra k(resp. $\mathfrak{a}$). On the compact group K the Haar measure dk is normalized such that the total measure is 1 (See [1] [2]).

Notations : We denote by

- $L^p(K\backslash G/K)$, $p \geq 1$, the space of functions f on G, bi-invariant under K, measurable and such that $\|f\|_p = (\int_G |f(x)|^p dx)^{1/p} < +\infty$. When G has rank one, we put : $f(a_t) = f[t]$, $a_t \in A$, $t \in \mathbb{R}$, the function f[t] is even, and then $L^p(K\backslash G/K)$ will be denoted $L^p(\mathbb{R}_+, A(t)dt)$, where $A(t) = 2^{2\rho}(\text{sh}t)^{2\alpha+1}(\text{ch}t)^{2\beta+1}$, with $\alpha = \frac{1}{2}(m_\alpha + m_{2\alpha} - 1)$, $\beta = \frac{1}{2}(m_{2\alpha} - 1)$, $\rho = \alpha + \beta + 1$. (See [4] p 14-16 and p 27).

- $L^p(\mathfrak{a}^*, |w|^{-1}|C(\lambda)|^{-2}d\lambda)^W$, $p \geq 1$, the space of functions f on $\mathfrak{a}^*$, W-invariant, mesurable and such that

$\|f\|_{L^p} = (\frac{1}{|w|}\int_{\mathfrak{a}^*}|f(\lambda)|^p|C(\lambda)|^{-2}\,d\lambda)^{1/p} < +\infty$, where C is the Harish-Chandra's C-function (See [1] p 175).

Definition I.1 : We said that a function g on G, mesurable, bi-invariant under K, is a wavelet on G if there is a constant Cg with the property that $o < C_g < +\infty$ and, for λ almost evywhere on $\mathfrak{a}^*$, $C_g = \int_o^\infty |\mathcal{F}(g)(a\lambda)|^2 \frac{da}{a}$. Where $\mathcal{F}$ is the spherical Fourier transform on G.

Example : We denote by β_t, $t > 0$, the function in the spherical Schwartz space $\mathcal{C}(K\backslash G/K)$ (See [1] p 252-257), such that

$$\forall\, \lambda \in \mathfrak{a}^*,\ \mathcal{F}(\beta_t)(\lambda) = \exp[-t(\|\lambda\|^2+\|\rho\|^2]$$

Then the function $g(x) = -\frac{d}{dt}\beta_t(x) - \|\rho\|^2\beta_t(x)$ is a wavelet on G which belongs to $\mathcal{C}(K\backslash G/K)$ and we have $C_g = \frac{e^{-t\|\rho\|^2}}{2t}$

If G has rank one the definition I.1 can be written more simply as follows : A function g on G measurable, bi-invariant under K, is a wavelet on G if the even function g[t] on $\mathbb{R}$ satisfies the condition $o < C_g = \int_o^\infty |\mathcal{F}(g)(a)|^2 \frac{da}{a} < +\infty$ (I.1), where $\mathcal{F}$ is the Jacobi transform (See [4] p 27)

Proposition I.1 : Let $g \neq 0$ be a function in $L^2(\mathbb{R}_+, A(t)dt)$ such that the function $\mathcal{F}(g)$ is continuous at 0 and there exists $\alpha > 0$ which satisfy : $\mathcal{F}(g)(\lambda) - \mathcal{F}(g)(0) = 0(\lambda^\alpha)$, as $\lambda \to 0$. Then the condition (I.1) is equivalent to $\mathcal{F}(g)(0) = 0$.

To define the dilated of a wavelet on G we need the following result on the Harish-Chandra's C-function.

Lemme I.1 : Let $\alpha \in \Sigma_o^+$. We have

$$k(a) = \sup_{\lambda \in \mathfrak{a}^*\setminus\{o\}} \frac{|C(\frac{\lambda}{a})|^{-2}}{|C(\lambda)|^{-2}} = \prod_{\alpha\in\Sigma_o^+} \begin{cases} \frac{1}{a^2}, & if\ a \geq 1 \\ \frac{1}{a^{m_\alpha+m_{2\alpha}}}, & if\ a < 1 \end{cases}$$

Theorem I.1 : Let g be a wavelet on G in $L^2(K\backslash G/K)$ and $a > 0$. Then the function $\lambda \to \mathcal{F}(g)(a\lambda)$ belongs to $L^2(\mathfrak{a}^*, |w|^{-1}|C(\lambda)|^{-2}d\lambda)^W$ and we have

$$\|\mathfrak{F}(g)(a)\|_{L^2} \leq (k(a)/a^{\ell})^{1/2}\|g\|_2$$

Moreover there exists a function $g_a \in L^2(K\backslash G/K)$ such that

$$\forall \lambda \in \mathfrak{a}^*, \mathfrak{F}(g_a)(\lambda) = \mathfrak{F}(g)(a\lambda)$$

and we have

$$\|g_a\|_2 \leq (k(a)/a^{\ell})^{1/2} \|g\|_2$$

Theorem I.2 : Let g be a wavelet on G in $\mathcal{D}(K\backslash G/K)$ (the space of C^∞-functions on G, bi-invariant under K and with compact support) (resp. $\mathcal{C}(K\backslash G/K)$) and g_a, $a > 0$, the function in $\mathcal{D}(K\backslash G/K)$ (resp $\mathcal{C}(K\backslash G/K)$) defined by

$$\forall \lambda \in \mathfrak{a}^*, \mathfrak{F}(g_a)(\lambda) = \mathfrak{F}(g)(a\lambda)$$

Then

$$\forall x \in G, g_a(x) = \mathbf{A}^{-1} \circ \mathcal{H}_a \circ \mathbf{A}(g)(x)$$

where $\mathbf{A}$ is the Abel transform on G (See [1] p 107, [2] p 450) and $\mathcal{H}_a$ the operator defined on $\mathfrak{a}$ by : $\mathcal{H}_a(f)(H) = a^{-\ell} f(\frac{H}{a})$

Definition I.2 : Let g be a wavelet on G in $\mathcal{D}(K\backslash G/K)$ (resp $\mathcal{C}(K\backslash G/K)$) (resp $L^2(K\backslash G/K)$) and g_a, $a > 0$, the function given by theorems I.1, I.2. We define the family of wavelets $g_{a,y}$, $(a,y) \in]0,+\infty[\; G$, on G by

$$\forall x \in G, g_{a,y}(x) = (k(a)/a^{\ell})^{-1/2} T_x(g_a)(y)$$

where T_x, $x \in G$, are generalized translation operators on G given by

$$T_x(f)(y) = \int_K f(xky)dk$$

Definition I.3 : Let g be a wavelet on G in $\mathcal{D}(K|G/K)$. We define the continuous wavelet transform on G for $f \in \mathcal{D}(K\backslash G/K)$ by

$$\Phi_g(f)(y) = \int_G f(x) \bar{g}_{a,y}(x)dx, \text{ for all } y \in G$$

Theorem I.3 : Let g be a wavelet on G in $(L^p \cap L^2)(K\backslash G/K)$, $p \in [1,2[$, such that for all $a > 0$, the function g_a is in $L^p(K\backslash G/K)$, $p \in [1,2[$. Then for all $f \in L^2(K\backslash G/K)$ we have

$$\|f\|_2^2 = \frac{1}{C_g} \int_0^\infty \int_G |\Phi_g(f)(a,y)|^2 \frac{k(a)}{a^{\ell+1}} da dy \text{ (Plancherel formula for } \Phi_g)$$

$$f(.) = \frac{1}{C_g} \int_0^\infty \int_G \Phi_g(f)(a,y)\, g_{a,y}(.) \frac{k(a)}{a^{\ell+1}} da\, dy \text{ (Inversion formula for } \Phi_g)$$

By theorem I.3 the transform Φ_g is an isometry of the Hilbert space $L^2(K\backslash G/K)$ into the Hilbert space L^2 $(]0,+\infty[\times G, \frac{k(a)}{a^{\ell+1} C_g} da dy)$ (the space of functions on $]0,+\infty[\times G$, bi-invariant under K with respect to the second variable, and square integrable on $]0,+\infty[\times G$ with respect to the measure $\frac{k(a)}{a^{\ell+1} C_g} da dy$). For the characterization of the image of Φ_g we interpret the vectors $g_{a,y}$, $(a,y) \in]0,+\infty[\times G$, as a set of coherent states in the Hilbert space $L^2(K\backslash G/K)$ (See [3] section I.2, [5] p 37 - 38).

<u>Theorem</u> I.4 : Let Φ_g be the continuous wavelet transform on G, with g a wavelet on G satisfying the assumptions of the theorem I.3. Let F be in L^2 $(]0,+\infty[\times G, \frac{k(a)}{a^{\ell+1} C_g} da dy)$. Then there exists a function $f \in L^2(K\backslash G/K)$ such that: $F = \Phi_g(f)$, if and only if

$$F(a,x) = \frac{1}{C_g} \int_0^\infty \int_G F(a',x') \left(\int_G g_{a',x'}(y) \bar{g}_{a,x}(y)\, dy\right) \frac{k(a')}{(a')^{\ell+1}}\, da'\, dx'$$

I give now an other inversion formula for the transform Φ_g.

<u>Theorem</u> I.5 : Let g be a wavelet on G in $L^2(K\backslash G/K)$. If f is a continuous and bounded function on G which belongs to $L^1(K\backslash G/K)$ and $\mathcal{F}(f) \in L^1(\mathfrak{a}^*, |w|^{-1} |C(\lambda)|^{-2} d\lambda)^W$. Then we have

$$f(x) = \frac{1}{C_g} \int_0^\infty \left(\int_G \Phi_g(f)(a,y)\, g_{a,y}(x) dy\right) \frac{k(a)}{a^{\ell+1}}\, da\,, \quad x \in G$$

<u>Theorem</u> I.6 : Let g be a wavelet on G in $\mathcal{D}(K\backslash G/K)$ (resp $\mathcal{S}(K\backslash G/K)$). Then for all f in $\mathcal{D}(K\backslash G/K)$ (resp $\mathcal{S}(K\backslash G/K)$) we have

$$\forall (a,y) \in]0,+\infty[\times G,\ \Phi_g(f)(a,y) = (k(a))^{-1/2} \mathcal{A}^{-1}[S_{\mathcal{A}(g)}(\mathcal{A}(f))(a,.)](y)$$

Where S is the classical continuous wavelet transform on $\mathbb{R}^\ell$ (See [5] p 27 - 45).

II. <u>Wavelets and continuous wavelet transform on the Cartan motion group</u> p.

Let G_0 denote the semidirect product K.p where p as a translation group is a normal subgroup and K acts on p by $Ad_G(K)$. We can view p as the quotient G_0/K which is a symmetric space of the Euclidean type. The group G_0 is often called the Cartan motion group associated with X = G/K. (See [2] p 467).

We prove for p the same results as for the group G, in particular we have

- The function k(a) is of the form

$$k(a) = \frac{1}{a^r}\ , \text{ with } r = \sum_{\alpha \in \Sigma^+} m_\alpha$$

-Let g be a wavelet on p in $L^2_K(p)$ (the space of functions f on p invariant under K, measurable and such that $\|f\|_2=(\int_p |f(y)|^2dy)^{1/2}<+\infty$).

Then the dilated function g_a, $a>0$, of g is given by : $g_a(x) = \frac{1}{a^{\ell+r}} g(\frac{x}{a})$,

and we have

$$\|g_a\|_2 = a^{-\frac{\ell+r}{2}} \|g\|_2$$

-We can take wavelets in $L^2_K(p)$ in the theorem which gives Plancherel and inversion formulas for the continuous wavelet transforms on p.

REFERENCES

[1] R. GANGOLLI - V.S.VARADARAJAN : Harmonic Analysis of spherical Functions on real reductive groups. Vol 101, Springer-Verlag Berlin, New-York, 1988.

[2] S. HELGASON : Groups and geometric Analysis, Integral Geometry, Invariant Differential operators and spherical functions. Academic Press, New-York, 1984.

[3] J. A. KLAUDER - B.S.SKAGERSTAM : A coherent state primer. In coherent states, J.R. Klauder and B.S.Skagerstam, eds., World Scientific, 1985, pp 1 - 115.

[4] T.H. KOORNWINDER : Jacobi functions and Analysis on noncompact semisimple Lie groups "Special functions : group theoretical aspects and applications" (R.A.Askey, T.H. Koornwinder and W. Schempp, eds.) Reidel, Dordrecht, 1984.

[5] T.H.KOORNWINDER : The continuous wavelet transform. Series in Approximations and Decompositions. Vol. 1. Wavelets : An elementary treatment of theory and Applications. Edited by T.H. Koornwinder, World Scientific, 1993, pp 27-48

Department of Mathematics, Faculty
of Sciences of Tunis, 1060 Tunis - TUNISIA

Received June 7. 1994

On a theorem of Carlitz-Von Staudt

*by Pieter Moree** (Princeton, USA)

Presented by M.R. Murty, F.R.S.C.

Abstract. Using some facts from the theory of primitive roots, a (corrected) generalized version of a result of Carlitz is proven. This result plays a crucial role in the analysis of certain diophantine equations first considered by Erdös and Moser.

Carlitz [C, Theorem 4] stated and, by combinatorial means, proved Theorem 1 below in the case $a = 0$ and $d = 1$. He falsely claimed the l.h.s. of (1) to be *always* an integer in case $r > 1$ and odd (it isn't: take any y with $y \equiv 2 \pmod 4$). He ascribes the result to Von Staudt, but doesn't give a reference. Since his proof, being a combination of elements from several previous results, is rather confusing, I didn't try to pinpoint the mistake in his reasoning. In this note I will give an elementary 'hands-on' proof of a generalization of his result, based only on some facts from the theory of primitive roots. The note will be ended with a discussion of applications of Theorem 1 in the analysis of Erdös-Moser type diophantine equations, that is diophantine equations similar to $1^k+2^k+\cdots+(x-1)^k = x^k$.

Let $S_r(a,d,y)$ denote the denominator of the l.h.s. of (1). For $S_r(0,1,y)$ we use the shorthand $S_r(y)$.

Theorem 1. *For arbitrary a, $d \geq 1$ with $(a,d) = 1$ we have*

$$\frac{a^r + (a+d)^r + \cdots + (a+d(y-1))^r}{y} \equiv \begin{cases} 0 \left(\bmod \frac{1}{2}\right), & r > 1 \text{ odd;} \\ -\sum_{p-1|r,\ p|y,\ p|d} \frac{1}{p} \pmod 1, & r \text{ even.} \end{cases} \tag{1}$$

* *Supported by the Netherlands Organization for Scientific Research* (NWO).

For the convenience of the reader we recall the proof of [MRU, Lemma 2]:

Lemma 1. *If p is odd and $p-1 \nmid r$, then $S_r(p^\lambda) \equiv 0 \pmod{p^\lambda}$ for every $\lambda \geq 1$.*

Proof. Notice that modulo p^λ, $S_r(p^\lambda)$ is unchanged by multiplication with g^r, where g is any primitive root modulo p (which exists for every odd prime). Since $p-1 \nmid r$ by assumption, $g^r \not\equiv 1 \pmod p$. Together with $p^\lambda | (g^r-1)S_r(p^\lambda)$ it follows from this that $S_r(p^\lambda) \equiv 0 \pmod{p^\lambda}$. ■

Lemma 2.

(i) *If $r > 1$ is odd and $\lambda \geq 2$, then $S_r(2^\lambda) \equiv 0 \pmod{2^\lambda}$.*

(ii) *If r is even and $\lambda \geq 1$, then $S_r(2^\lambda) \equiv 2^{\lambda-1} \pmod{2^\lambda}$.*

Proof. (i) Use that $j^r \equiv -(2^\lambda - j)^r \pmod{2^\lambda}$.

(ii) For $\lambda = 1$ and 2 the assertion is clear, so assume $\lambda \geq 3$. From the theory of primitive roots it follows that modulo 2^λ every odd number is of the form 5^j or $3 \cdot 5^j$ with $0 \leq j < 2^{\lambda-2}$. We find that

$$S_r(2^\lambda) \equiv [1+3^r]\frac{5^{2^{\lambda-2}r}-1}{5^r-1} + 2^r S_r(2^{\lambda-1}) \pmod{2^\lambda}. \tag{2}$$

We have $5^r = 1 + k2^{\mathrm{ord}_2(r)+2}$ with k odd.

$$5^{r2^{\lambda-2}} \equiv 1 + k2^{\lambda+\mathrm{ord}_2(r)} \pmod{2^{\lambda+\mathrm{ord}_2(r)+1}}$$

and $1+3^r \equiv 2 \pmod 4$. So the first term on the r.h.s. of (2) is congruent to $2^{\lambda-1} \pmod{2^\lambda}$. On using induction the proof becomes complete. ■

Lemma 3. *If p is odd and $p-1|r$, then $S_r(p^\lambda) \equiv -p^{\lambda-1} \pmod{p^\lambda}$ for every $\lambda \geq 1$.*

Proof. By the theory of primitive roots there exists a number g which is a primitive root modulo p^λ for all $\lambda \geq 1$. Since $p-1|r$, $g^r = 1 + kp^f$ for some k not divisible by p and $f \geq 1$. From this it follows that

$$g^{\frac{\varphi(p^\lambda)r}{p-1}} - 1 = g^{p^{\lambda-1}r} - 1 \equiv kp^{\lambda+f-1} \pmod{p^{\lambda+f}}$$

and so $g^{\varphi(p^\lambda)r} - 1 \equiv 1 + k(p-1)p^{\lambda+f-1} \pmod{p^{\lambda+f}}$. Thus we have

$$\frac{g^{\varphi(p^\lambda)r}-1}{g^r-1} \equiv -p^{\lambda-1} \pmod{p^\lambda}. \tag{3}$$

Clearly the lemma holds for $\lambda = 1$. On using (3), the fact that $r \geq 2$, induction on λ and the relation

$$S_r(p^\lambda) \equiv \frac{g^{\varphi(p^\lambda)r}-1}{g^r-1} + p^r S_r(p^{\lambda-1}) \pmod{p^\lambda}, \tag{4}$$

the lemma follows. The congruence (4) follows on expressing the sum of the terms in $S_r(p^\lambda)$ coprime to p^λ as a geometric series in g^r. ∎

Proof of Theorem 1. It suffices to show that for every prime q dividing y with $e(q)$ such that $q^{e(q)} \| y$, we have

$$S_r(a,d,y) \equiv \begin{cases} 0 \pmod{q^{e(q)}}, & r > 1 \text{ odd}; \\ -\sum_{\substack{p-1|r,\ p|y \\ p|d}} \frac{y}{p} \pmod{q^{e(q)}}, & r \text{ even}. \end{cases} \tag{5}$$

In case $r > 1$ odd and $q = 2$ it suffices to show that $S_r(a,d,y) \equiv 0 \pmod{2^{e(2)-1}}$.

Case 1. $q \nmid d$ and $q - 1 \nmid r$. We have to show that for $r > 1$, $S_r(a,d,y) \equiv 0 \pmod{q^{e(q)}}$. Notice that q must be odd. We have $S_r(a,d,q^{e(q)}) \equiv S_r(q^{e(q)}) \pmod{q^{e(q)}}$, so $S_r(a,d,y) \equiv \frac{y}{q^{e(q)}} S_r(q^{e(q)}) \pmod{q^{e(q)}}$. Therefore by Lemma 1, $S_r(a,d,y) \equiv 0 \pmod{q^{e(q)}}$.

Case 2. $q \nmid d$ and $q - 1 | r$. Again we have $S_r(a,d,y) \equiv \frac{y}{q^{e(q)}} S_r(q^{e(q)}) \pmod{q^{e(q)}}$. In case $q = 2$ we have to show that $S_r(2^{e(2)}) \equiv 0 \pmod{2^{e(2)-1}}$. This follows by Lemma 2. In the case q is odd, r is even and we have to show that $S_r(q^{e(q)}) \equiv -q^{e(q)-1} \pmod{q^{e(q)}}$. The validity of this congruence follows by Lemma 3.

Notice that for $d = 1$ this completes the proof. In case $d > 1$, we have to investigate the remaining case where $q|d$. We have

$$S_r(a,d,y) = \sum_{j=0}^{k} \binom{k}{j} a^{k-j} S_j(y) d^j,$$

where $S_0(y) = y$. It suffices to show that $S_j(y)d^j \equiv 0 (\text{mod } q^{e(q)})$ for every $j \geq 1$. We have $S_1(y) = \frac{y(y-1)}{2}d \equiv 0(\text{mod } q^{e(q)})$. For $j > 1$ it easily follows from (5) for $d = 1$, which we already have shown to hold true, that $S_j(y)d^j \equiv 0(\text{mod } q^{e(q)})$. This concludes the proof.

Remark. A slight refinement of the analysis shows that in case $r > 1$ and odd, the l.h.s. of (1) is an integer if and only if d is even or $4|y$.

In [M] I considered the equation $1^k + \cdots + (x-1)^k = ax^k$ and various other equations of this form. I showed that the former equation has no integer solutions (a, x, k) with $k > 1$ and $x < \max\{10^{10^6}, a10^{22}\}$. This equation was probably first considered by Krzystzofek [K]. The special case $a = 1$ was considered by Moser [Mos]. Erdös conjectures that in this case there are no solutions with $k > 1$. In the proof of the above inequality for x, Theorem 1 in the case $a = 0$ and $d = 1$ plays a crucial role. To give some feeling for the way Theorem 1 is used, I will prove the identity $\text{ord}_2(x-3) = \text{ord}_2(k) + 3$, that holds in case $a = 1$. In [MRU, Lemma 12] a more complicated proof making use of Bernoulli polynomials, is given. So let x, k with $k > 1$ satisfy $1^k + \cdots + (x-1)^k = x^k$. Then $1^k + \cdots + (x-4)^k = x^k - (x-1)^k - (x-2)^k - (x-3)^k$. On applying Theorem 1 with $a = 0, d = 1, y = x - 4$ we obtain:

$$\frac{3^k - 1 - 2^k}{x-3} \equiv - \sum_{\substack{p-1|k \\ p|x-3}} \frac{1}{p} \ (\text{mod } 1).$$

We use the fact that k must be even, ≥ 6 and x must be odd. Notice that the 2-order of the r.h.s. is -1. The 2-order of the l.h.s. must also be -1, i.e. $\text{ord}_2(x-3) = \text{ord}_2(3^k - 1 - 2^k) + 1$. On using induction and the fact that $k \geq 6$, it is easily seen that

$$\text{ord}_2(x-3) = \text{ord}_2(3^k - 1 - 2^k) + 1 = \text{ord}_2(k) + 3.$$

REFERENCES

[C] L. Carlitz, The Staudt-Clausen theorem, *Math. Mag.* 34 (1961), 131-146.

[K] B. Krzystzofek, The equation $1^n + \cdots + m^n = (m+1)^n k$ (Polish), *Wyz. Szkol. Ped. w. Katowicach-Zeszyty Nauk. Sekc. Mat.* 5 (1966), 47-54.

[M] P. Moree, On the integer solutions of $1^k + 2^k + \cdots + (x-1)^k = ax^k$, Leiden University Report W94-05, March 1994, submitted.

[MRU] P. Moree, H.J.J. te Riele and J. Urbanowicz, Divisibility properties of integers x and k satisfying $1^k + 2^k + \cdots + (x-1)^k = x^k$, to appear in Mathematics of Computation.

[Mos] L. Moser, On the diophantine equation $1^n + 2^n + \cdots + (m-1)^n = m^n$, *Scripta Math.* 19 (1953), 84-88.

Freesiastraat 13
2651 XL Berkel en Rodenrijs
The Netherlands
email: moree@rulfcl.leidenuniv.nl

Received July 6, 1994

Mailing Addresses

R.V. Chacon	Department of Mathematics University of British Columbia Vancouver, B.C., Canada, V6T 1Y4
A.T. Fomenko	Department of Topology and Geometry Moscow State University 119899 Moscow, Russia
B. Gilligan	Department of Mathematics and Statistics University of Regina Regina, Saskatchewan, Canada, S4S 0A2
A. Lyzzaik	Department of Mathematics American University of Beirut Beirut, Lebanon
P. Moree	Freesiastraat 13 2651 XL Berkel en Rodenrijs The Netherlands
K. Trimeche	Department of Mathematics Faculty of Sciences of Tunis 1060 Tunis - Tunisia