

ROSTOCKER

MATHEMATISCHES KOLLOQUIUM

Heft 47

UNIVERSITÄT



ROSTOCK

HERAUSGEBER: Der Rektor der Universität Rostock

WISSENSCHAFTLICHE LEITUNG: Prof.Dr. Hans-Wolfgang Stolle
(Sprecher des Fachbereiches Mathematik)

Dr. Werner Plischke

REDAKTIONELLE BEARBEITUNG: Dr. Werner Plischke

HERSTELLUNG DER DRUCKVORLAGE: S. Dittmer
W. Hartmann
H. Schubert

REDAKTIONSSCHLUS: 15. Dezember 1993

Universität Rostock
Mathematisch-Naturwissenschaftliche Fakultät
FB Mathematik
D - 18051 Rostock

Das **Rostocker Mathematische Kolloquium** erscheint dreimal im Jahr und ist im Rahmen des Schriftentausches über die Universität Rostock, Universitätsbibliothek, Tauschstelle, zu beziehen.

ZITAT KURZTITEL: Rostock. Math. Kolloq. (1994) 47

ISSN 0138-3248

© Universität Rostock, Presse- und Informationsstelle, Wissenschaftspublizistik, 18051 Rostock

BEZUGSMÖGLICHKEITEN: Universität Rostock, Presse- und Informationsstelle, Wissenschaftspublizistik
A.-Einstein-Straße 23, 18051 Rostock, ☎ (0381) 4 405 520; FAX (0381) 4 405 434
Universität Rostock, Universitätsbibliothek, Schriftentausch, 18051 Rostock

DRUCK: Universitätsdruckerei 43/94

01000

https://doi.org/10.18453/rosdok_id00003879

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 47

EBERHARD OELJEKLAUS	<i>Meromorphic function fields of Albanese bundles</i>	3
WOLFGANG BANNUSCHER; GUNTER TIEDT	<i>On a theorem of Deaconescu</i>	23
W. B. VASANTHA KANDASAMY	<i>Supermodular subgroups in finite groups</i>	27
FRIEDRICH LIESE; ANDREAS WIENKE	<i>Asymptotische Normalität von Maximum - Likelihood - Schätzern in inhomogenen Wienerprozessen</i>	31
WALTRAUT KAHLE; FRIEDRICH LIESE	<i>Consistency and asymptotic normality of minimum contrast estimation in renewal processes</i>	49
JENS MEYER	<i>Epidemiemodelle und Iterierte Funktionssysteme</i>	63
BURKHARD LENZE	<i>Design and realization of a family of optimal windows</i>	78
WOLFGANG ENGEL	<i>Bildungseinrichtungen für mathematisch begabte Schüler in Rostock</i>	91

UNIVERSITÄT ROSTOCK

FACHBEREICH MATHEMATIK

1994

EBERHARD OELJEKLAUS

Meromorphic function fields of Albanese bundles¹

Introduction

Let X be a (connected) compact complex manifold and $\text{Aut } X$ the complex Lie group of biholomorphic transformations of X . The Albanese torus $A(X)$ of X is a bimeromorphic invariant of X , and the holomorphic Albanese map $\alpha : X \rightarrow A(X)$ often yields important information about the holomorphic or meromorphic structure of X . For example, if X is kählerian and if $c_1(X) \geq 0$ or $\text{Aut } X$ has an open orbit on X , then $\alpha : X \rightarrow A(X)$ is a holomorphic fiber bundle and much is known about the topological and holomorphic structure of X (cf. [18], [19], [20], [13], [1], [14]). In particular, if $\text{Aut } X$ operates transitively on X , then X splits biholomorphically into a product $A(X) \times F$ with a projective rational manifold F ([3]).

In general not much can be said about the structure of the general fiber F of the Albanese map, even if $\alpha : X \rightarrow A(X)$ is holomorphically locally trivial. Therefore we assume that $\text{Aut } X$ acts transitively on $A(X)$ via α and call $\alpha : X \rightarrow A(X)$ a *homogeneous* bundle. We want to find sufficient criteria for a bimeromorphic splitting of X into a product $A(X) \times F$. Simple examples show that such a splitting is in general impossible if X is not projective algebraic.

Lieberman [14] and Fujiki [6] analyzed the structure of $\text{Aut}^0 X$ for a compact Kähler manifold X . Using results of Barlet and Douady about cycle spaces they showed in particular that every $\text{Aut}^0 X$ -orbit on X is Zariski-open in its closure and that in case of a homogeneous Albanese bundle $\text{Aut}^0 X$ is an extension of the kernel of $\alpha_* : \text{Aut}^0 X \rightarrow \text{Aut}^0 A(X)$ by the center of $\text{Aut}^0 X$. If X is projective algebraic it follows immediately from their work together with results of Rosenlicht [22] and Matsumura [15] that a finite unramified covering \tilde{X} of X splits birationally into a product $A(\tilde{X}) \times F$.

Therefore the problem of birational splitting of the (homogeneous) Albanese bundle of a projective algebraic manifold X can be reduced to the study of bundles with finite abelian structure group.

¹Ausgearbeitete und erweiterte Fassung eines Vortrags, gehalten auf dem 3. Bremen-Rostocker Kolloquium, Rostock, 22. und 23. Mai 1992

Let K be a rational function field over \mathbb{C} and $\Gamma \subset \text{Aut}_{\mathbb{C}}K$ a subgroup. If (z_1, \dots, z_n) is a transcendence basis of $K = \mathbb{C}(z_1, \dots, z_n)$ over \mathbb{C} with $\Gamma(\mathbb{C}(z_1, \dots, z_i)) = \mathbb{C}(z_1, \dots, z_i)$, $1 \leq i \leq n$, then we say that Γ *fixes a flag* in K . In [24] it is shown that the invariant fields of finite groups Γ of odd order with the flag-fixing property fulfill strong rationality conditions. We apply these rationality properties to prove our main result (Theorem 15):

Let X be a projective algebraic manifold with homogeneous Albanese bundle with projective rational fiber F . There is an abelian variety T , a projective unirational manifold \tilde{F} and a finite subgroup $\Gamma \subset T$, embedded into $\text{Aut } \tilde{F}$, such that X is birationally equivalent to $(T \times_{\Gamma} \tilde{F}) \times \mathbb{P}^k$. If Γ is of odd order and fixes a flag in the meromorphic function field $\mathfrak{M}(\tilde{F})$ of \tilde{F} , then X splits birationally into a product $A(X) \times \mathbb{P}^r$.

The splitting property of the Albanese bundle depends strongly on the fact that there are enough meromorphic functions on $A(X)$. If $A(X)$ is not an abelian variety then in general it is impossible to get a bimeromorphic splitting by going to a finite unramified covering, even if the fiber of the bundle is a rational curve:

Proposition 14 *Let X be a compact Kähler manifold and $G \subset \text{Aut } X$ a connected complex subgroup with an open orbit Ω on X . Assume that Ω cannot be holomorphically fibered by compact complex tori. The following statements are equivalent:*

- i) *There is a finite unramified covering \tilde{X} of X which splits bimeromorphically into a product $A(\tilde{X}) \times F$.*
- ii) *$A(X)$ is an abelian variety.*
- iii) *X is projective algebraic.*

Using cycle space methods, Campana has proved uniqueness properties for the algebraic reduction of a compact complex manifold X . Using his results (see [4]) we can prove:

Proposition 19 *Let X be an almost homogeneous compact complex manifold. There is a projective algebraic almost homogeneous manifold \tilde{X} with $\mathfrak{M}(\tilde{X}) = \mathfrak{M}(X)$. The Albanese torus $A(\tilde{X})$ is the algebraic reduction of $A(X)$.*

This note is organized as follows: In section 1 we discuss the main structural properties of homogeneous kählerian Albanese bundles. In section 2 we give a straightforward and elementary proof of the complex version of a splitting theorem of Rosenlicht [22] and Matsushima [15]. This direct approach yields the detailed information about the orbits of an algebraic action of a connected solvable linear algebraic group which we need for the proof of theorem 15. In particular it enables us in our special situation to transform rational group actions of algebraic groups into regular ones via a birational mapping without using the structure theorems of A. Weil [26] in their general form.

1 G-induced Albanese bundles

For a (connected) compact complex manifold X the Albanese torus $A(X)$ and the holomorphic Albanese map $\alpha : X \rightarrow A(X)$ are characterized by the universality property that every holomorphic map $\pi : X \rightarrow T$ into a compact complex torus T induces a unique holomorphic map $\pi_A : A(X) \rightarrow T$ with $\pi_A \circ \alpha = \pi$. Consequently α is $\text{Aut} X$ -equivariant and defines a holomorphic homomorphism

$$\alpha_* : \text{Aut}^0 X \rightarrow \text{Aut}^0 A(X) = A(X), \quad \alpha \circ g = \alpha_*(g) \circ \alpha,$$

where $\text{Aut}^0 A(X)$ is identified with $A(X)$ by fixing a basepoint $0 \in A(X)$. The Albanese torus $A(X)$ is a bimeromorphic invariant of X with $\dim_{\mathbb{C}} A(X) \leq \frac{1}{2} b_1(X)$ and $\dim_{\mathbb{C}} A(X) = \frac{1}{2} b_1(X)$ if X is kählerian. The Albanese torus of a projective algebraic manifold is an abelian variety (for details see [2]).

In general not much can be said about the structure of the general fiber F of the Albanese map. Even if X is projective algebraic and α is a locally trivial holomorphic fibration it may happen that F is a complex torus with $b_1(F) = b_1(X)$, e.g. if X is a hyperelliptic surface. If $X \neq A(X)$ is a compact complex solvmanifold, then X is a holomorphic fiber bundle over $A(X)$ and F is again a compact complex solvmanifold; in particular $b_1(F) \neq 0$.

If X is projective algebraic, the group $I := \ker \alpha_*$ is a linear algebraic group: By [2], Th. I.3.2 the canonical operation of I on $H^1(X, \mathcal{O}^*)$ is trivial. Take a very ample line bundle ζ on X and an embedding $\varphi : X \rightarrow \mathbb{P}_m$ defined by a basis of $\Gamma(X, \mathcal{O}(\zeta))$. Then the I -action on $\varphi(X)$ is induced by a linear action I on \mathbb{P}_m . Therefore the algebraic closure \bar{I} of I in $\text{PGL}(m, \mathbb{C})$ stabilizes $\varphi(X)$ and \bar{I} is a linear algebraic group.

Let G be a connected complex subgroup of $\text{Aut} X$. We call $\alpha : X \rightarrow A(X)$ G -induced, if $\alpha_*(G) = A(X)$. In this case $F := \alpha^{-1}(0)$ is connected and α is a locally trivial holomorphic fibration, namely the quotient of $G \times F$ by the following action ρ of $G \cap \ker \alpha_*$ on $G \times F$:

$$\rho(h)(g, z) := (gh^{-1}, h(z)), \quad h \in G \cap \ker \alpha_*, g \in G, z \in F.$$

Remarks 1) If G has an open orbit on X , then $\alpha : X \rightarrow A(X)$ is G -induced. Nevertheless it may happen that $b_1(F) \neq 0$, e.g. if X is a non-kählerian compact complex solvmanifold. But if X is kählerian, then $H^1(F, \mathcal{O}) = H^2(F, \mathcal{O}) = 0$ ([21]). In particular F is projective algebraic and $b_1(F) = 0$.

2) In [1] the structure of $\text{Aut}^0 X$ was analyzed for a compact Kähler manifold X with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ and projective algebraic fiber F with $b_1(F) = 0$. It was shown that $\ker \alpha_*$ is a linear algebraic group and that the canonical homomorphism $\ker \alpha_* \rightarrow \text{Aut} F$ is injective.

3) D. Lieberman [14] and A. Fujiki [6] independently studied the structure of $\text{Aut}^0 X$ for arbitrary Kähler manifolds X . They used cycle space methods and got very general results. In particular they showed that every $\text{Aut}^0 X$ -orbit is Zariski-open in its Zariski-closure. If $\alpha : X \rightarrow A(X)$ is G -induced, then $\ker \alpha_*$ is a linear algebraic group. ■

For G -induced Albanese bundles the situation $b_1(F) \neq 0$ can only occur when $\ker \alpha_*$ has infinitely many components.

Lemma 1 *Let X be a compact complex manifold with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ with typical fiber F . If $I := \ker \alpha_*$ has only finitely many components, then $b_1(F) = 0$.*

Proof: We set $G := \text{Aut}^0 X$. The Albanese map $\beta : F \rightarrow A(F)$ yields a representation $\beta_* : I \rightarrow \text{Aut } A(F)$. Since $\beta_*(I^0) \subset \text{Aut}^0 A(F) = A(F)$, the group $G \times A(F)$ operates holomorphically and transitively on the compact manifold $\tilde{Y} := G \times_{I^0} A(F)$, which is a holomorphic $A(F)$ -bundle over the complex torus G/I^0 . It follows that \tilde{Y} is complex parallelisable and a complex Lie group, since the isotropy group L of $[(e, 0)] \in \tilde{Y}$ is connected:

$$L = \{(h, \beta_*(h^{-1})) \in I^0 \times A(F)\} = L^0.$$

Therefore \tilde{Y} and its finite quotient $Y := G \times_I A(F)$ are complex tori. The canonical G -equivariant holomorphic map $X = G \times_I F \rightarrow Y$ together with the universality of $\alpha : X \rightarrow A(X)$ yield $Y = A(X)$, i.e. $A(F) = \{0\}$. ■

From the fact that $\ker \alpha_*$ is a linear algebraic group it follows that $\text{Aut}^0 X$ is an extension of $\ker \alpha_*$ by the center $C(\text{Aut}^0 X)$ of $\text{Aut}^0 X$. More precisely:

Lemma 2 *Let X be a compact complex manifold and G a connected complex subgroup of $\text{Aut } X$. If the Albanese map $\alpha : X \rightarrow A(X)$ is G -induced and if $G \cap \ker \alpha_*$ is a linear algebraic group, then G is an extension of $G \cap \ker \alpha_*$ by the center $C(G)$ of G ; in other words $G = (G \cap \ker \alpha_*) \cdot C(G)$. In particular α is $C(G)$ -induced.*

Proof: Let \mathfrak{i} denote the Lie algebra of $I := G \cap \ker \alpha_*$ and define $\text{Ad}_I G$ to be the image of $\text{Ad } G$ under the restriction map $\text{Ad } G \rightarrow \text{Aut } \mathfrak{i}$, $\text{Ad } G \mapsto \text{Ad } G|_{\mathfrak{i}}$. Let $C_G(I^0)$ and $C_I(I^0)$ denote the centralizer of I^0 in G and I , respectively.

The adjoint representation of the linear algebraic group I is regular, hence $\text{Ad } I$ is a linear algebraic normal subgroup of $\text{Ad}_I G$ and of the algebraic closure of $\text{Ad}_I G$ in $\text{GL}(\mathfrak{i})$. Therefore $\text{Ad}_I G / \text{Ad } I$ is holomorphically separable and $A(X) = G/I \rightarrow \text{Ad}_I G / \text{Ad } I$ has to be constant. Then $\text{Ad}_I G = \text{Ad } I$, $G/I = C_G(I^0)/C_I(I^0) = C_G(I^0)/I \cap C_G(I^0)$, and $G = I \cdot C_G(I^0) = I^0 \cdot C_G^0(I^0)$.

Since $G/I = C_G^0(I^0)/C_I(I^0) \cap C_G^0(I^0)$ is compact and since $C_I(I^0)$ is linear algebraic, the group $C_G^0(I^0)/C_I^0(I^0)$ is a compact complex torus. The relation $C_I^0(I^0) \subset C(C_G^0(I^0))$ implies that $C_G^0(I^0)$ is abelian and $C_G^0(I^0) \subset C^0(G)$. ■

If the Albanese map $\alpha : X \rightarrow A(X)$ is G -induced with typical fiber $F = \mathbf{P}_1$, then X is kählerian and again we can apply Remark 2) together with Lemma 2. This leads to

Proposition 3 *Let G be a connected complex Lie group and $H \subset I \subset G$ closed complex subgroups with $T := G/I$ a compact complex torus and $I/H \in \{\mathbf{C}^*, \mathbf{C}, \mathbf{P}_1\}$. If G is non-abelian, then the bundle $G/H \rightarrow G/I$ is holomorphically trivial.*

Proof: There is a natural G -equivariant kählerian compactification X of G/H , namely a holomorphic \mathbb{P}_1 -bundle $\alpha : X \rightarrow T = A(X)$, and we may assume that $G = \text{Aut}^0 X$ and $I = \ker \alpha_*$. By Lemma 2 we have $G = I \cdot C^0(G)$ and since G is non-abelian it follows that $\dim I \geq 2$. Therefore $I = \text{PGL}(2, \mathbb{C})$ or I is a parabolic subgroup of $\text{PGL}(2, \mathbb{C})$. In particular $C(G) \cap I \subset C(I) = \{e\}$. This implies $G = I \times C(G) = I \times C^0(G) = I \times A(X)$, $G/H = I/H \times A(X)$ and $X = \mathbb{P}_1 \times A(X)$. ■

Corollary *If $\alpha : X \rightarrow A(X)$ is a G -induced \mathbb{P}_1 -bundle and if $\text{Aut}^0 X$ is non-abelian, then X splits biholomorphically into a product $\mathbb{P}_1 \times A(X)$.* ■

The condition for $\text{Aut}^0 X$ to be non-abelian is really essential. Otherwise we meet counterexamples already for projective algebraic surfaces:

Example 1: The embedding $\mathbb{C} \rightarrow \mathbb{C}^* \times \mathbb{C}^*$, $z \rightarrow (e^z, e^{\tau z})$, $\tau \in \mathbb{C}$, $\text{Im } \tau > 0$, realizes \mathbb{C} as a closed complex subgroup of $\mathbb{C}^* \times \mathbb{C}^*$, the quotient $T := \mathbb{C}^* \times \mathbb{C}^* / \mathbb{C}$ being an elliptic curve. The canonical compactification of the \mathbb{C} -principal bundle $\mathbb{C}^* \times \mathbb{C}^* \rightarrow T$ is a projective algebraic \mathbb{P}_1 -bundle X over $T = A(X)$. No finite unramified covering \tilde{X} of X splits biholomorphically into a product $\mathbb{P}_1 \times A(\tilde{X})$, since $\mathbb{C}^* \times \mathbb{C}^* \not\subset A(\tilde{X}) \times \text{PGL}(2, \mathbb{C}) = \text{Aut}^0(A(\tilde{X}) \times \mathbb{P}_1)$. Note that $\text{Aut}^0 X = \mathbb{C}^* \times \mathbb{C}^*$ has an open orbit on \tilde{X} . ■

This example is also interesting under the following aspect: For any projective algebraic manifold X the group $\text{Aut}^0 X$ carries the structure of an algebraic group (see [17], [16]), and for a G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ the group $\ker \alpha_*$ is linear algebraic. In the above example $\text{Aut}^0 X = \mathbb{C}^* \times \mathbb{C}^*$ is a linear algebraic group, but $\ker \alpha_*$ is not a linear algebraic subgroup of $\text{Aut}^0 X$. In particular $\text{Aut}^0 X$ is not a linear algebraic transformation group of X .

Remember that a (connected) compact complex manifold X is called *homogeneous*, if $\text{Aut } X$ operates transitively on X . We call X *almost homogeneous*, if $\text{Aut } X$ has an open orbit on X . Meanwhile it is a classical theorem that the Albanese bundle $\alpha : X \rightarrow A(X)$ of a compact homogeneous Kähler manifold splits biholomorphically into a product $A(X) \times F$, where F is projective rational. It was proven by Matsushima ([18]) for the case of reductive groups and in general by Borel-Remmert in [3]. This splitting is no longer true for almost homogeneous manifolds, as the above example shows. For almost homogeneous kählerian manifolds X the fiber F of the Albanese bundle is unirational, but in general not rational. A finite unramified covering \tilde{X} of X splits topologically into a product $A(\tilde{X}) \times F$, but $\alpha : X \rightarrow A(X)$ is in general not topologically trivial ([1]).

We are interested in the birational structure of a G -induced Albanese bundle and in the following we will study this problem. It seems to be still an open problem whether the Albanese bundle of an almost homogeneous projective algebraic manifold X splits birationally into a product $A(X) \times F$.

As is known and will be seen (Corollary to Proposition 11) a suitable finite unramified covering \tilde{X} of a projective algebraic G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ splits birationally into a product $A(\tilde{X}) \times F$. This is not true for compact Kähler manifolds, even if they are almost homogeneous:

Example 2: Let Γ_{2n-1} be a discrete subgroup of \mathbb{C}^n of rang $2n - 1$. The group $\mathbb{C}^n/\Gamma_{2n-1}$ can be fibered as a holomorphic \mathbb{C}^* -principal bundle over a compact complex torus T . The canonical compactification X of this bundle is an almost homogeneous \mathbb{P}_1 -bundle over $T = A(X)$. For $n \geq 3$ the group Γ_{2n-1} can be chosen such that $\mathbb{C}^n/\Gamma_{2n-1}$ has no non-constant meromorphic functions. Then no finite unramified covering \tilde{X} of X splits birationally into a product $A(\tilde{X}) \times \mathbb{P}_1$. ■

For a compact complex manifold Y we define $a(Y)$ to be the transcendence degree over \mathbb{C} of the field of meromorphic functions on Y and $d(Y) := \dim_{\mathbb{C}} Y - a(Y)$. A compact almost homogeneous Kähler manifold X is positive algebraic if and only if $d(A(X)) = 0$ (Proposition 5).

Lemma 4 *Let X be a compact Kähler manifold and $G \subset \text{Aut } X$ a connected complex subgroup with an open orbit $\Omega = G/H$ in X . Assume that a finite unramified covering \tilde{X} of X splits bimeromorphically into a product $F \times A(\tilde{X})$. There is a closed complex subgroup I of G containing H and a holomorphic fibration $\Omega = G/H \rightarrow G/I$ with $T := I/H$ a $d(A(X))$ -dimensional compact complex torus.*

Proof: For $x \in X$ define A_x to be the intersection of all closed complex hypersurfaces in X containing x . Then $g(A_x) = A_{g(x)}$ for all $g \in \text{Aut } X$ and $A_x \cap \Omega = A_y \cap \Omega$ or $A_x \cap A_y \cap \Omega = \emptyset$ for all $x, y \in \Omega$. We fix $x \in \Omega$. The group $I := \{g \in G \mid g(A_x) = A_x\}$ operates transitively on $A_x \cap \Omega$. Let A be the irreducible component of A_x containing $A_x \cap \Omega$. Then $a(X) = \text{codim } A$ since G operates transitively on Ω (see [5]).

Now \tilde{X} splits and the fiber F of the Albanese bundle $\alpha : X \rightarrow A(X)$ is projective algebraic, therefore

$$a(X) = \dim F + a(A(X)), \quad \dim(A \cap F) = 0,$$

and

$$\dim \alpha(A) = \dim A = \dim X - a(X) = \dim A(X) - a(A(X)) = d(A(X)).$$

Since I operates with an open orbit on A and on $\alpha(A)$ and since all I -orbits in $A(X)$ have the same dimension, the group I operates already transitively on $\alpha(A)$, and $\alpha(A)$ is a compact complex torus. The assumption $A \not\subseteq \Omega$ leads to the contradiction

$$d(A(X)) = \dim \alpha(A) \leq \dim(A \cap (X \setminus \Omega)) \leq \dim A - 1 \leq d(A(X)) - 1.$$

This completes the proof. ■

Example 3: Embed \mathbb{C}^2 as a closed complex subgroup into $G := (\mathbb{C}^*)^4$ such that the compact complex torus $T := G/\mathbb{C}^2$ is not projective algebraic. The canonical G -equivariant compactification X of G as a holomorphic \mathbb{P}_2 -bundle over $T = A(X)$ is an almost homogeneous Kähler manifold. Since $\Omega := G$ is Stein, no finite unramified covering \tilde{X} of X splits meromorphically into a product $\mathbb{P}_2 \times A(\tilde{X})$.

A holomorphic \mathbb{P}_n -bundle X over a projective algebraic base Y splits birationally into a product if it is induced by a rang $(n + 1)$ -vector bundle over Y , because every such vector

bundle is an algebraic bundle ([23]). In particular X splits birationally into a product $\mathbb{P}_n \times Y$, if $H^2(Y, \mathcal{O}^*) = 0$, e.g. if Y is a compact Riemann surface.

Every G -induced kählerian Albanese bundle $\alpha : X \rightarrow A(X)$ over an abelian variety $A(X)$ with projective algebraic fiber F is a subbundle of a holomorphic \mathbb{P}_n -bundle $\tilde{X} \rightarrow A(X)$.

Proposition 5 *Let X be a compact Kähler manifold with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ where the fiber F is projective algebraic. The following statements are equivalent:*

- i) $A(X)$ is an abelian variety.
- ii) The Albanese bundle $X \rightarrow A(X)$ is a subbundle of a holomorphic \mathbb{P}_n -bundle $\tilde{X} \rightarrow A(X)$.
- iii) X is projective algebraic.

Proof: It remains to verify the implication i) \Rightarrow ii). Since $I = \ker \alpha_*$ has only finitely many components and since $b_1(F) = 0$ there is a proper embedding of F into \mathbb{P}_n such that the I -action on F is induced by a linear I -action on \mathbb{P}_n . Choosing n minimal we see that $X = G \times_I F$ is a closed submanifold of $G \times_I \mathbb{P}_n$, which is a holomorphic \mathbb{P}_n -bundle over $A(X)$ and therefore projective algebraic ([2]). ■

2 Reduction to finite structure groups

Studying the birational structure of the Albanese bundle we are lead to meromorphic G -actions. Let G be a complex Lie group and Y an irreducible compact complex space. A meromorphic map $\psi : G \times Y \rightarrow Y$ is called a *meromorphic G -action* on Y if $\psi(g, -) : Y \rightarrow Y$ defines a bimeromorphic map of Y for every $g \in G$ with $\psi(e, -) = \text{id}_Y$ and $\psi(gh, y) = \psi(g, \psi(h, y))$ whenever $\psi(gh, y)$, $\psi(h, y)$ and $\psi(g, \psi(h, y))$ are defined. If G is an algebraic group, Y a projective algebraic variety and ψ a rational (resp. regular) map, we call ψ a *rational* (resp. *regular*) G -action on Y . In the above situation G operates as a group of \mathbb{C} -automorphisms on the meromorphic function field $\mathfrak{M}(Y)$ of Y with invariant field $\mathfrak{M}(Y)^G$.

If B is a connected solvable linear algebraic group with a rational action on a projective algebraic variety Y , then, by the cross-section-theorem of Rosenlicht ([22]) generalized by Matsumura ([15]), the field $\mathfrak{M}(Y)$ is a purely transcendental extension of $\mathfrak{M}(Y)^B$, say $\mathfrak{M}(Y) = \mathfrak{M}(Y)^B(z_1, \dots, z_k)$. Let Y_B be a projective variety with $(Y_B) = \mathfrak{M}(Y)^B$. The above result means that Y is birationally equivalent to $Y_B \times \mathbb{P}^k$.

Assume now that Y is projective rational. Then Y_B is stably rational but need not be rational. There is a rational manifold Z on which B acts regularly and effectively with an open orbit. Take a rational manifold Y of the form $Y = W \times Z$ where W is unirational but not rational. The B -action on Z leads to an example where $\mathfrak{M}(Y)^B = \mathfrak{M}(W)$ is not a rational function field. Note (see [25]) that $\mathfrak{M}(Y)^B$ is a rational function field if and only if B fixes a flag in $\mathfrak{M}(Y)$, i.e. there is a transcendence basis (z_1, \dots, z_n) of $\mathfrak{M}(Y) = \mathbb{C}(z_1, \dots, z_n)$ with $B(\mathbb{C}(z_1, \dots, z_i)) = \mathbb{C}(z_1, \dots, z_i)$ for $i \leq i \leq n$.

The purpose of this section is a straightforward and rather elementary proof of a special version of the Rosenlicht-theorem. This direct approach yields detailed information about the orbits of an algebraic action of a connected solvable linear algebraic group which we need later. In particular it allows us in our situation to transform rational group actions into regular ones via a birational mapping without using the general structure theorems of A. Weil ([26]).

For the sake of completeness we start with the regular-section-theorem for algebraic actions of solvable linear groups on affine varieties over \mathbb{C} .

Proposition 6 *Let Z be a normal irreducible affine algebraic variety and B a connected solvable linear algebraic group operating regularly and effectively on Z .*

- 1) *There is a B -invariant Zariski-open affine subvariety U of Z , a Zariski-closed subvariety V of U and a biregular map*

$$\sigma : \mathbb{C}^m \times (\mathbb{C}^*)^n \times W \rightarrow U, \quad m, n \geq 0,$$

with $\sigma(\mathbb{C}^m \times (\mathbb{C}^)^n, w) = B(w)$ for all $w \in W$.*

- 2) *Every B -orbit in Z is biregularly equivalent to a product of the form $\mathbb{C}^r \times (\mathbb{C}^*)^s$.*

The main step for the proof consists in analyzing regular \mathbb{C} - and \mathbb{C}^* -actions:

Lemma 7 *Let Z and B be as in Proposition 6. For every onedimensional connected normal algebraic subgroup N of B there is a B -invariant Zariski-open affine subvariety U_1 of Z and a Zariski-closed subvariety W_1 of U such that*

$$\sigma_1 : N \times W_1 \rightarrow U_1, \quad \sigma_1(n, w) := n(w),$$

is biregular. With $p : N \times W_1 \rightarrow W_1$ the canonical surjection, the map

$$\psi : B/N \times W_1 \rightarrow W_1, \quad \psi(gN, w) := p \circ \sigma^{-1}(g(w)),$$

defines a regular B/N -action on W_1 such that $N(\psi(B/N)(w)) = B(w)$ for all $w \in W_1$.

Proof of Proposition 6: By induction on $\dim_{\mathbb{C}} B$. Let $N, U_1, W_1, \sigma_1 : N \times W_1 \rightarrow U_1$ and $\psi : B/N \times W_1 \rightarrow W_1$ be as in Lemma 7 and let N_1/N be the kernel of ineffectivity of the ψ -action of B/N on W_1 . By induction there is a biregular map

$$\sigma_2 : \mathbb{C}^m \times (\mathbb{C}^*)^n \times W \rightarrow U_2$$

with U_2 an affine algebraic Zariski-open subvariety of W_1 and W a Zariski-closed subvariety of U_2 and

$$B/N_1(w) = \sigma_2(\mathbb{C}^m \times (\mathbb{C}^*)^n, w) \quad \text{for all } w \in W.$$

The composition

$$N \times \mathbb{C}^m \times (\mathbb{C}^*)^n \times W \xrightarrow{\text{id} \times \sigma_2} N \times U_2 \xrightarrow{\sigma_1} \sigma_1(N \times U_2) =: U$$

defines a biregular map $\sigma := \sigma_1 \circ (\text{id} \times \sigma_2)$ from $N \times \mathbb{C}^m \times (\mathbb{C}^*)^n \times W$ onto U . For $w \in W$ we have $\sigma(N \times \mathbb{C}^m \times (\mathbb{C}^*)^n, w) = N(\psi(B/N)(w)) = B(w)$. This completes the proof of 1). Statement 2) follows immediately from 1) by induction on $\dim Z$. \blacksquare

Proof of Lemma 7: We embed Z as a closed affine subvariety into \mathbf{C}^n such that the B -action on Z is induced by a linear action of B on \mathbf{C}^n (see [12]) and choose n minimal. Identifying \mathbf{C}^n with $\{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid u_n \neq 0\}$ we study the connected solvable complex groups $\tilde{N} \subset \tilde{B} \subset \mathrm{SL}(n+1, \mathbf{C})$ with $\tilde{B}/\Gamma \cap \tilde{B} = B$, $\tilde{N}/\Gamma \cap \tilde{N} = N$, where $\Gamma \simeq \mathbf{Z}_{n+1}$ denotes the center of $\mathrm{SL}(n+1, \mathbf{C})$.

By the flag theorem of Lie we may assume that \tilde{B} is contained in the subgroup of upper triangular matrices of $\mathrm{SL}(n+1, \mathbf{C})$ and handle the two possibilities $\tilde{N} = \mathbf{C}^*$ and $\tilde{N} = \mathbf{C}$ separately.

a) $\tilde{N} = \mathbf{C}^*$

Since \tilde{N} is a connected normal subgroup of \tilde{B} consisting only of semisimple elements, it is contained in the center of \tilde{B} (see [10]). All elements of \tilde{N} are diagonal matrices because the maximal tori in \tilde{B} are conjugate to each other. Assume that

$$\tilde{N} = \{\mathrm{Diag}(\alpha^{k_0}, \dots, \alpha^{k_n}) \in \mathrm{SL}(n+1, \mathbf{C}) \mid \alpha \in \mathbf{C}^*, k_j \in \mathbf{Z}, \sum_{j=0}^n k_j = 0\}.$$

Define

$$\kappa(j) := \max\{i \in \{0, 1, \dots, n\} \mid k_i = k_j\} \text{ for } j \in \{0, 1, \dots, n\}$$

and assume that there are exactly $r+1$ elements $\kappa_0, \dots, \kappa_r$ in the set $\{\kappa(j) \mid j \in \{0, 1, \dots, n\}\}$ for some $r \geq 1$. For $d := \mathrm{gcd}(k_{\kappa_1} - k_{\kappa_0}, \dots, k_{\kappa_r} - k_{\kappa_0})$ there is a canonical isomorphism

$$\varphi: \mathbf{C}^*/\mathbf{Z}_d \rightarrow N, \varphi([\alpha]) := \mathrm{Diag}(\alpha^{k_0}, \dots, \alpha^{k_n}) \bmod \Gamma.$$

Take $a_\nu \in \mathbf{Z}$ with $d = \sum_{\rho=1}^r a_\rho (k_{\kappa_\rho} - k_{\kappa_0})$ and choose $t_j, t'_j \in \mathbf{N}$ with

$$t_0 - t'_0 = \sum_{\rho=1}^r a_\rho, \quad t'_\rho - t_\rho = a_\rho, \quad \rho \in \mathbf{N}_r.$$

In particular

$$\sum_{j=0}^r t_j = \sum_{j=0}^r t'_j, \quad d = \sum_{j=0}^r k_{\kappa_j} (t'_j - t_j).$$

Define

$$f(x_0, \dots, x_n) := \prod_{j=0}^r x_{\kappa_j}^{t_j} + \prod_{j=0}^r x_{\kappa_j}^{t'_j} \in C[x_0, \dots, x_n]$$

and

$$H := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid f(u_0, \dots, u_n) = 0\} \cap Z.$$

Since \tilde{N} centralizes \tilde{B} , the hyperplanes

$$A_{\kappa_j} := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid u_{\kappa_j} = 0\}, \quad 0 \leq j \leq r,$$

are \tilde{B} -invariant by construction, and it follows that the Zariski-open affine subvariety

$$U_1 := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid \prod_{j=0}^r u_{\kappa_j} \neq 0\} \cap Z$$

of Z is B -invariant. For $z = (u_0 : \dots : u_n) \in U$ we get

$$\begin{aligned} f(\varphi([\alpha])(z)) &= \alpha^{\sum_{j=0}^r k_{\kappa_j} t_j} \left(\prod_{j=0}^r u_{\kappa_j}^{t_j} \right) + \alpha^{\sum_{j=0}^r k_{\kappa_j} t'_j} \left(\prod_{j=0}^r u_{\kappa_j}^{t'_j} \right) \\ &= \alpha^{\sum_{j=0}^r k_{\kappa_j} t_j} \left(\prod_{j=0}^r u_{\kappa_j}^{t_j} + \alpha^d \cdot \prod_{j=0}^r u_{\kappa_j}^{t'_j} \right). \end{aligned}$$

This shows that there is exactly one $[\alpha] \in \mathbb{C}^*/\mathbb{Z}_d \cong N$ with $\varphi([\alpha])(z) \in H \cap U_1 =: W_1$. It follows that $\sigma_1 : N \times W_1 \rightarrow U_1$, $\sigma(g, w) := g(w)$, is a bijective regular open map. Since U_1 is normal, σ_1 is biregular.

b) $\tilde{N} = \mathbb{C}$

In this case $\tilde{N} \cap \Gamma = \{e\}$ and we may assume $n \geq 2$. There is a nilpotent upper triangular matrix $A \in \mathbb{C}^{(n+1, n+1)}$ of the form

$$\begin{pmatrix} 0 & a_{01} & \dots & a_{0k} & \dots & \dots & a_{0n} \\ 0 & 0 & & & & & \vdots \\ \vdots & \vdots & \ddots & a_{k-1, k} & \dots & \dots & a_{k-1, n} \\ & & & 0 & r_{k+1} & \dots & r_n \\ & & & & 0 & \dots & 0 \\ & & & & & \ddots & \vdots \\ 0 & 0 & \dots & & & & 0 \end{pmatrix},$$

$0 \leq k \leq n-1$, $(r_{k+1}, \dots, r_n) \neq (0, \dots, 0)$, with $\tilde{N} = \{e^{\lambda A} \mid \lambda \in \mathbb{C}\}$. The linear hyperplane

$$\tilde{L} := \{z = (z_0, \dots, z_n) \in \mathbb{C}^{n+1} \mid (Az)_k = \sum_{j=1}^{n-k} r_{k+j} z_{k+j} = 0\}$$

is \tilde{B} -invariant: Take $P = (p_{ij}) \in \tilde{B}$, $0 \leq i, j \leq n$. There exists $\alpha \in \mathbb{C}^*$ with

$$e^{PAP^{-1}} = P e^A P^{-1} = e^{\alpha A}.$$

Since PAP^{-1} and αA are upper triangular nilpotent matrices it follows that $PAP^{-1} = \alpha A$. For $PA = (c_{ij})$, $0 \leq i, j \leq n$, we get

$$c_{kj} = \begin{cases} 0 & 0 \leq j \leq k \\ p_{kk} r_j & k+1 \leq j \leq n \end{cases}$$

and

$$(PAz)_k = \sum_{j=k+1}^n p_{kk} r_j z_j = p_{kk} \sum_{j=1}^{n-k} r_{k+j} z_{k+j} = p_{kk} (Az)_k.$$

This means that

$$(A(Pz))_k = (\alpha^{-1} PAz)_k = \alpha^{-1} p_{kk} (Az)_k$$

and that \tilde{L} is \tilde{B} -invariant. In particular, the hyperplane

$$L := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid \sum_{j=1}^{n-k} r_{k+j} u_{k+j} = 0\}$$

is B -invariant. With

$$U_1 := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid u_n \cdot \sum_{j=1}^{n-k} r_{k+j} u_{k+j} \neq 0\} \cap Z,$$

$$H := \{(u_0 : \dots : u_n) \in \mathbb{P}_n \mid u_k = 0\} \cap Z,$$

and $u := (u_0 \dots : u_n) \in U_1$ we calculate $e^{\lambda A}(u) := (w_0 : \dots : w_n)$ and find

$$w_k = u_k + \sum_{j=1}^{n-k} \lambda r_{k+j} u_{k+j} = u_k + \lambda \cdot \sum_{j=1}^{n-k} r_{k+j} u_{k+j}.$$

Since $\sum_{j=1}^{n-k} r_{k+j} u_{k+j} \neq 0$, there is exactly one $\lambda \in \mathbb{C}$ with $e^{\lambda A}(u) \in H \cap U_1 =: W_1$. Just as above we conclude that $\sigma_1 : N \times W_1 \rightarrow U_1$ is biregular. The second statement of the Lemma is trivial. ■

The splitting behaviour described in Proposition 6 leads immediately to a birational splitting of a normal projective algebraic variety X . Note that the kernel L of the canonical representation of $\text{Aut} X$ on $H^1(X, \mathcal{O}^*)$ is a linear algebraic group with $L^0 = (\ker \alpha)^0 \subset \ker \alpha \subset L$ if X is a projective algebraic manifold with Albanese map α .

Proposition 8 *Let X be a normal projective algebraic variety and B a connected solvable linear algebraic subgroup of L . There is a closed irreducible subvariety Y of X and a birational map*

$$\tau : \mathbb{P}_k \times Y \rightarrow X, \quad k = \max\{\dim B(x) \mid x \in X\}.$$

There is a Zariski-open subvariety V of Y with $\tau(\mathbb{P}_k \times \{v\}) = \overline{B(v)}$ for every $v \in V$. In particular $\mathfrak{M}(Y) = \mathfrak{M}(X)^B$.

Proof: We embed X into \mathbb{P}_n such that the B -action on X is induced by a linear action of B on \mathbb{P}_n . By the flag theorem of Lie B stabilizes a hyperplane section D of X and the Zariski-open affine submanifold $Z := X \setminus D \neq \emptyset$. Using the notation of Proposition 6 we get a Zariski-open B -invariant affine subvariety U of Z , a closed subvariety W of U and a biregular map

$$\sigma : \mathbb{C}^m \times (\mathbb{C}^*)^n \times W \rightarrow U, \quad B(w) = \sigma(\mathbb{C}^m \times (\mathbb{C}^*)^n, w).$$

With $Y := \overline{W}$ (Zariski-closure of W in X) and $\hat{\sigma} : \mathbb{P}_m \times \mathbb{P}_n \times Y \rightarrow X$ the birational extension of σ we get the desired $\tau : \mathbb{P}_{m+n} \times Y \rightarrow \mathbb{P}_n \times \mathbb{P}_n \times Y \xrightarrow{\hat{\sigma}} X$. ■

We want to study the subvariety Y more precisely. For a subgroup $B \subset \text{Aut} X$ we define $N(B)$ to be the normalizer of B in $\text{Aut} X$.

Proposition 9 *Let X be a projective algebraic manifold with Albanese map $\alpha : X \rightarrow A(X)$ and B a connected solvable linear algebraic subgroup of $\ker \alpha_*$ and G a connected complex subgroup of $(N(B))^0$ containing B . Let $\tau : \mathbb{P}_k \times Y \rightarrow X$ be a birational splitting as in Proposition 8; in particular $\alpha(Y) = \alpha(X)$.*

1) *The meromorphic map*

$\psi : G/B \times Y \rightarrow Y$, $\psi(gB, y) := p \circ \tau^{-1} \circ g(y)$, where $p : \mathbb{P}_k \times Y \rightarrow Y$ is the canonical surjection, defines a meromorphic G/B -action on Y and is compatible with the holomorphic G/B -action on $A(X)$:

$$\alpha(\psi(gB, y)) = \alpha_*(g) \circ \alpha(y) \quad \text{for all } g \in G, y \in Y.$$

2) *There is a Zariski-open subset U of X such that $\alpha^{-1}(\alpha(u)) =: F_u$ is birationally equivalent to $\mathbb{P}_k \times (Y \cap F_u)$ for all $u \in U$.*

Proof: 1) Because of the special properties of Y (see Proposition 8) a short calculation shows that ψ defines in fact a meromorphic G/B -action on Y . Since α is G -equivariant and since $B \subset \ker \alpha_*$, this action is compatible with the G/B -action on $A(X)$.

2) With U, W, σ as in Proposition 8 the map

$$\sigma|_{F_u} : \mathbb{C}^m \times (\mathbb{C}^*)^n \times (W \cap F_u) \rightarrow U \cap F_u$$

is biregular for every $u \in U$ and induces a birational map

$$\mathbb{P}_k \times (Y \times F_u) \rightarrow \mathbb{P}_m \times \mathbb{P}_n \times (Y \times F_u) \rightarrow F_u.$$

■

A meromorphic G -action on a complex space Y may be quite unpleasant for obvious reasons. We want to transform the meromorphic G/B -action on Y into a holomorphic G/B -action on a birationally equivalent projective algebraic variety \tilde{Y} . In our special situation this is possible.

Lemma 10 *Let T be a compact complex torus and $\psi : T \times Y \rightarrow Y$ a meromorphic T -action on an irreducible compact complex variety Y . Let Γ be a finite subgroup of T and $p : T \rightarrow T/\Gamma$ the canonical surjection. Assume that $\alpha : Y \rightarrow T/\Gamma$ is a holomorphic map with*

$$p(t)(\alpha(y)) = \alpha \circ \psi(t, y), \quad t \in T, y \in Y,$$

whenever $\psi(t, y)$ is defined.

There exists a normal compact complex variety \tilde{F} and a holomorphic Γ -action on \tilde{F} with the following properties:

1) $\tilde{Y} := T \times_{\Gamma} \tilde{F}$ is bimeromorphically equivalent to Y .

2) \tilde{F} is bimeromorphically equivalent to the general α -fiber.

3) If Y is projective algebraic, then \tilde{F} can be chosen to be projective algebraic, too.

Proof: Left translation of $\Gamma = \{\gamma_1 := e, \gamma_2, \dots, \gamma_n\}$ on itself induces an embedding ϕ of Γ into the permutation group of n elements. Let $Y := Y \times \dots \times Y$ (k factors) and define $\hat{Y} \subset Y_n$ to be the graph of the meromorphic map

$$(\gamma_2, \dots, \gamma_n) : Y \rightarrow Y_{n-1}, y \mapsto (\gamma_2(y), \dots, \gamma_n(y)),$$

whenever $\gamma_2(y), \dots, \gamma_n(y)$ are defined. Then Γ operates holomorphically on \hat{Y} by permutation:

$$\gamma(y_1, \dots, y_n) := (y_{\phi(\gamma)(1)}, \dots, y_{\phi(\gamma)(n)}), \quad \gamma \in \Gamma.$$

Since the group T is abelian, the Γ -action can be extended to a meromorphic T -action $T \times \hat{Y} \rightarrow \hat{T}$ with $t(y_1, \dots, y_n) := (t(y_1), \dots, t(y_n))$, $t \in T$, whenever $t(y_1), \dots, t(y_n)$ are defined.

The projection of Y_n onto the first factor induces a holomorphic and bimeromorphic map of \hat{Y} onto Y . Replacing Y by \hat{Y} we may assume that Γ operates holomorphically on Y .

Let \tilde{F} be the normalization of a general fiber of $\alpha : Y \rightarrow T/\Gamma$. The meromorphic T -action ψ induces a generically surjective meromorphic map $\tilde{\psi} : T \times \tilde{F} \rightarrow Y$ and T operates holomorphically on the quotient $\hat{Y} := T \times_{\Gamma} \tilde{F}$, $\gamma(t, z) := (t - \gamma, \gamma(z))$, $\gamma \in \Gamma$, $t \in T$, $z \in \tilde{F}$. The meromorphic map

$$\hat{Y} \rightarrow Y, [t, z] \mapsto \tilde{\psi}(t, z) = t(z),$$

is bimeromorphic.

Remark: A general theorem of A. Weil ([26]) says, that every rational G -action on a projective algebraic variety Y can be transformed into a regular one: There is a birational map $\varphi : Y \rightarrow \hat{Y}$ onto a Moishezon variety \hat{Y} such that the induced G -action on \hat{Y} is regular. ■

Proposition 11 *Let X be a projective algebraic manifold with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ with typical fiber F such that $I := G \cap \ker \alpha_*$ is a linear algebraic group (e.g. $G = \text{Aut}^0 X$). Let C be the identity component of the center of G , $T := C/C \cap I^0$ and $\Gamma := C \cap I/C \cap I^0$. Let B be a Borel subgroup of I and $k := \max\{\dim Bx | x \in X\}$. There is a normal projective algebraic variety \tilde{F} and a holomorphic Γ -action on \tilde{F} with the following properties:*

1) *There is a birational map*

$$\tau : \mathbf{P}_k \times (T \times_{\Gamma} \tilde{F}) \rightarrow X.$$

2) *The typical fiber F of α is birationally equivalent to $\mathbf{P}_k \times \tilde{F}$ and $\mathfrak{M}(\tilde{F}) = \mathfrak{M}(F)^B$.*

Proof: Note that $G = CI$ by Lemma 2. The center $C(I) = C \cap I$ of I is a linear algebraic group and every Borel subgroup B of I^0 contains $C(I^0) = C \cap I^0 = C \cap B$. Applying Proposition 9 to the group CB we get a birational splitting $\tau : \mathbf{P}_k \times Y \rightarrow X$ and a meromorphic action of the abelian variety $T = CB/B = C/C \cap B$ on Y . Now Proposition 11 follows from Lemma 10 in connection with statement 2) of Proposition 9. ■

Corollary *Let X be a projective algebraic manifold with G -induced bundle $\alpha : X \rightarrow A(X)$ with typical fiber F .*

- 1) *A suitable finite unramified covering \tilde{X} of X splits birationally into a product $F \times A(\tilde{X})$.*
- 2) *X is birationally equivalent to a projective algebraic manifold $T \times_{\Gamma} F$, where T is an abelian variety and $\Gamma \subset T$ a finite subgroup with $T/\Gamma = A(X)$.*

Summing up we have a sufficient criterion for the birational splitting of Albanese bundles in terms of linear algebraic groups:

Proposition 12 *Let X be a projective algebraic manifold with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ and typical fiber F . Assume that at least one of the following conditions is fulfilled for $I := \ker \alpha_* \subset \text{Aut } F$, $C := C^0(\text{Aut}^0 X)$, and the center $C(I) = C \cap I$ of I :*

- 1) $C(I) \subset I^0$ (i.e. $C(I) = C(I^0)$)
- 2) $C(I) \subset \text{Aut}^0 F$ and $C(I)$ contains a regular semisimple element of $\text{Aut}^0 F$.
- 3) $C(I)$ consists only of semisimple elements and is contained in a Borel subgroup B of $\text{Aut} F$.
- 4) A Borel subgroup of $\text{Aut } X$ operates with an open orbit on X .

Then X splits birationally into a product $A(X) \times F$.

Proof: 1) By Proposition 11 there is a birational splitting $X \rightarrow \mathbb{P}_k \times A(X) \times \tilde{F}$, and $\mathbb{P}_k \times \tilde{F}$ is birationally equivalent to F .

2) The centralizer in $\text{Aut}^0 F$ of a regular semisimple element is a maximal torus S in $\text{Aut}^0 F$. Therefore $C \cap I \subset S \subset I^0$ since S is abelian.

3) Every subgroup of B consisting only of semisimple elements is contained in a maximal torus S of B . As above we get $C \cap I \subset S \subset I^0$.

4) Every Borel subgroup B of I has an open orbit on F . By Proposition 11 it follows that X is birationally equivalent to $\mathbb{P}_k \times A(X)$. ■

Corollary *Let X be an almost homogeneous projective algebraic manifold with $\dim X \leq 2 + \dim A(X)$. Then $X = A(X)$ or X is birationally equivalent to $\mathbb{P}_k \times A(X)$.*

Proof: A Borel subgroup of $\text{Aut} X$ operates with an open orbit on X . ■

Remark: Note that in general the group $C \cap I$ is not contained in a Borel subgroup of $\text{Aut } F$, even if $C \cap I$ is finite and $\text{Aut } F$ is connected. Take an abelian variety T and $\Gamma := \mathbf{Z}_2 \times \mathbf{Z}_2 \subset T$ and identify Γ with the subgroup $\{z \mapsto \pm \frac{1}{z}\}$ of $\text{Aut } \mathbf{P}_1$. Then $X := T \times_{\Gamma} \mathbf{P}_1$ is a holomorphic \mathbf{P}_1 -bundle over $A(X) = T/\Gamma$ with $\text{Aut}^0 X = T$ and $I = C \cap I = \Gamma$, which is not contained in a Borel subgroup of $\text{Aut } \mathbf{P}_1$. ■

We know that every holomorphic \mathbf{P}_1 -bundle over a compact Riemann surface Y splits birationally into a product $\mathbf{P}_1 \times Y$, but for G -induced Albanese bundles with onedimensional fiber we have only the following result:

Proposition 13 *Let $\alpha : X \rightarrow A(X)$ be a holomorphic and G -induced \mathbf{P}_1 -bundle over the abelian variety $A(X)$. Then X is birationally equivalent to $\mathbf{P}_1 \times A(X)$ or $\text{Aut}^0 X$ is an abelian variety with $\text{Aut}^0 X / (\mathbf{Z}_2 \times \mathbf{Z}_2) = A(X)$ and $\dim X \geq 3$.*

Proof: If $\dim \ker \alpha_* \geq 1$ then the Corollary of Proposition 12 applies. If $\ker \alpha_*$ is finite, then $\text{Aut}^0 X$ is an abelian variety and $\ker \alpha_*$ equals its centralizer in $\text{Aut } \mathbf{P}_1$. But then $\ker \alpha_* = \mathbf{Z}_2 \times \mathbf{Z}_2$. ■

Finally we generalize the result of Lemma 4 about the splitting of compact almost homogeneous Kähler manifolds.

Proposition 14 *Let X be a compact Kähler manifold and $G \subset \text{Aut } X$ a connected complex subgroup with an open orbit Ω on X . Assume that Ω cannot be holomorphically fibered by compact complex tori. The following statements are equivalent:*

- i) *There is a finite unramified covering \tilde{X} of X which splits bimeromorphically into a product $A(\tilde{X}) \times F$.*
- ii) *$A(X)$ is an abelian variety.*
- iii) *X is projective algebraic.*

Proof: i) \Rightarrow ii): From Lemma 4 we get $d(A(X)) = 0$, hence $A(X)$ is an abelian variety.

ii) \Leftrightarrow iii): Proposition 5.

iii) \Rightarrow i): Corollary of Proposition 11. ■

3 Flag-fixing structure groups

Let us assume that we have a projective algebraic manifold X with G -induced Albanese bundle $\alpha : X \rightarrow A(X)$ with projective rational fiber F . For $G := \text{Aut}^0 X$, $I := \ker \alpha_*$ and C the identity component of the center of G we know that X is birationally equivalent to a holomorphic fiber bundle $(T \times_{\Gamma} \tilde{F}) \times \mathbf{P}_k$, where $T = C/C \cap I^0$ is an abelian variety and $\Gamma := C \cap I/C \cap I^0$ is a finite subgroup of T (Proposition 11). Therefore the problem of the birational splitting of the Albanese bundle of X reduces to the problem of the birational splitting of the holomorphic bundle $T \times_{\Gamma} \tilde{F}$ into a product $A(X) \times \tilde{F}$. Let B be a Borel subgroup of I . We assume that B fixes a flag in $\mathfrak{M}(F)$, i.e. that $\mathfrak{M}(F)^B = \mathfrak{M}(\tilde{F})$ is a rational function field.

Theorem 15 *If Γ is of odd order and fixes a flag in $\mathfrak{M}(\bar{F})$, then X splits birationally into a product $A(X) \times \mathbb{P}_r$.*

Let $K = \mathbb{C}(z_1, \dots, z_n)$ be a rational function field with transcendence basis (z_1, \dots, z_n) and Γ a finite subgroup of $\text{Aut}_{\mathbb{C}}K$ with $K_i := \mathbb{C}(z_1, \dots, z_i) = \Gamma(K_i)$, $1 \leq i \leq n$. In general the invariant field K^Γ of Γ is not a rational function field even if Γ is cyclic. There is an example (see [24]) for $n = 3$ and an involution $\sigma \in \text{Aut}_{\mathbb{C}}K$ with $\sigma(z_1) = z_1$, $\sigma(z_2) = -z_2$, where K^σ is not a rational function field. This example is based on results of Artin, Mumford and Deligne about unirational hypersurfaces in \mathbb{P}_4 which are not rational. But if Γ is of odd order, then K^Γ is always rational. This is a consequence of the following result which was proved in [24]:

If Γ is of odd order, then K_{i+1}^Γ is a purely transcendental extension of K_i^Γ , $0 \leq i \leq n-1$, $K_0 := \mathbb{C}$.

This result yields the induction step for the proof of the following splitting theorem, which implies Theorem 15 by the above remarks.

Theorem 16 *Let T be an abelian variety and $\Gamma \subset T$ a finite subgroup of odd order operating holomorphically on a k -dimensional projective rational manifold F . If Γ fixes a flag in $\mathfrak{M}(F)$, then $X := T \times_{\Gamma} F$ is birationally equivalent to $A(X) \times \mathbb{P}_k$.*

Let Γ and $K = \mathbb{C}(z_1, \dots, z_n)$ be as above. In the following we assume that Γ is of odd order.

Lemma 17 *The group Γ fixes a flag (x_1, \dots, x_n) in K with the property*

$$\gamma(x_n) = \alpha(\gamma)x_n, \quad \alpha(\gamma) \in \mathbb{C}(x_1, \dots, x_{n-1})$$

for all $\gamma \in \Gamma$.

For $\alpha \in \mathbb{C}^*$ define $g_\alpha \in \text{Aut}_{\mathbb{C}}K$ by $g_\alpha(x_i) := x_i + \delta_{in}(\alpha - 1)x_i$, $1 \leq i \leq n$. The group $A := \{g_\alpha \mid \alpha \in \mathbb{C}^*\}$ centralizes Γ and $K^A = \mathbb{C}(x_1, \dots, x_{n-1})$.

Proof: Let $r := \min\{i \in \mathbb{N}_m \mid \Gamma \rightarrow \text{Aut}_{\mathbb{C}}K_i \text{ is injective}\}$ and $\tilde{\gamma} \in \Gamma$ with $\tilde{\gamma}|_{K_r} \neq \text{id}$, $\tilde{\gamma}|_{K_{r-1}} = \text{id}$. We may assume that $\tilde{\gamma}(z_r) = \beta z_r$ where β is a root of unity with $\text{ord } \beta \geq 3$. For $\gamma \in \Gamma$ we get

$$\gamma(z_r) = \frac{az_r + b}{cz_r + d}, \quad a, b, c, d \in K_{r-1}, \quad ad - bc \neq 0,$$

and the relation $\tilde{\gamma} \circ \gamma = \gamma \circ \tilde{\gamma}$ yields

$$\frac{a\beta z_r + b}{c\beta z_r + d} = \beta \frac{az_r + b}{cz_r + d}.$$

Then $ac = bc = 0$, since $\beta^2 \neq 1$. In particular $b = c = 0$. By the above quoted result of [24] the invariant field K^Γ of Γ is a purely transcendental extension of K_r^Γ , say $K^\Gamma = K_r^\Gamma(u_{r+1}, \dots, u_n)$. Then

$$\begin{aligned} & [K_r(u_{r+1}, \dots, u_n) : K_r^\Gamma(u_{r+1}, \dots, u_n)] = \\ & [K_r(u_{r+1}, \dots, u_n) : K_r^\Gamma] \geq [K_r : K_r^\Gamma] = \text{ord } \Gamma = [K : K^\Gamma] \end{aligned}$$

and therefore $K = K_r(u_{r+1}, \dots, u_n) = \mathbb{C}(z_1, \dots, z_{r-1}, u_{r+1}, \dots, u_n, z_r)$.
 Define $x_j := z_j$, $1 \leq j \leq r-1$, $x_k := u_{k+1}$, $r+1 \leq k \leq n-1$, $x_n := z_r$. ■

Proposition 18 *Let F be a projective rational manifold and $\Gamma \subset \text{Aut } F$ a finite abelian group of odd order, fixing a flag in $\mathfrak{M}(F)$. There is a projective rational manifold \tilde{F} and a birational map $\psi : F \rightarrow \tilde{F}$ such that $\tilde{\Gamma} := \psi\Gamma\psi^{-1}$ is contained in a onedimensional abelian reductive subgroup R of $\text{Aut } \tilde{F}$ and fixes a flag in the rational function field $\mathfrak{M}(\tilde{F})^{R^0}$.*

Proof: With $K := \mathfrak{M}(F)$ we choose (x_1, \dots, x_r) and A as in Lemma 17 and get a birational map

$$\tau : F \rightarrow \mathbb{P}_1 \times \dots \times \mathbb{P}_1 =: \hat{F}, \quad y \mapsto (x_1(y), \dots, x_n(y)).$$

Then $\hat{R} := \tau A \Gamma \tau^{-1}$ is a group of birational transformations of \hat{F} . Note that $\hat{R}^0 = \tau A \tau^{-1} = \mathbb{C}^*$ operates holomorphically on \hat{F} . The $\hat{R} := \tau A \Gamma \tau^{-1}$ is a group of birational transformations of \hat{F} . Note that $\hat{R}^0 = \tau A \tau^{-1} = \mathbb{C}^*$ operates holomorphically on \hat{F} . The finite abelian group $\hat{\Gamma} := \tau \Gamma \tau^{-1} = \{\gamma_1 = \text{id}, \gamma_2, \dots, \gamma_r\}$ defines a rational map

$$\Lambda : \hat{F} \rightarrow \hat{F} \times \dots \times \hat{F}, \quad u \mapsto (\gamma_2(u), \dots, \gamma_r(u)),$$

and the canonical holomorphic and birational map

$$\pi : \text{graph } \Lambda \rightarrow \hat{F}, \quad (u_1, \dots, u_r) \mapsto u_1,$$

induces a π -equivariant holomorphic action of \hat{R} on $\text{graph } \Lambda$ with the property

$$g(u_1, \dots, u_r) = (g(u_1), \dots, g(u_r)), \quad g \in \hat{R}, \quad (u_1, \dots, u_r) \in \text{graph } \Lambda,$$

whenever $g(u_j)$ is defined for $1 \leq j \leq r$. With respect to this action \hat{R} can be identified with a complex subgroup $\pi^{-1} \hat{R} \pi \subset \text{Aut}(\text{graph } \Lambda)$. Let $\tilde{F} \xrightarrow{\rho} \text{graph } \Lambda$ be a desingularization of $\text{graph } \Lambda$ which is $\pi^{-1} \hat{R} \pi$ -equivariant. Then $\psi := \rho^{-1} \circ \pi^{-1} \circ \tau$ and $R := \psi(A \Gamma) \psi^{-1}$ have the desired properties. ■

Finally we prove Theorem 16 by induction on $k = \dim F$:

Using the notation of Lemma 17 we know that X is birationally equivalent to $\tilde{X} := T \times_{\Gamma} \tilde{F}$, where Γ is identified with $\psi \Gamma \psi^{-1} \in \text{Aut } \tilde{F}$ and fixes a flag in $\mathfrak{M}(\tilde{F})^{R^0}$. The group $\tilde{G} := T \times R^0$ operates holomorphically on \tilde{X} and by Proposition 11 there is a birational map $\tilde{X} \rightarrow (T \times_{\Gamma} \hat{Y}) \times \mathbb{P}_1$, where \hat{Y} is a normal projective rational variety with $\mathfrak{M} = (\tilde{F})^{R^0}$. Substituting \hat{Y} by a Γ -equivariant projective rational desingularization Y we may apply the induction hypothesis to $T \times_{\Gamma} Y \simeq T \times_{\Gamma} \hat{Y}$. This completes the proof. ■

Proposition 19 *Let X be an almost homogeneous compact complex manifold. There is a projective algebraic almost homogeneous manifold \hat{X} with $\mathfrak{M}(\hat{X}) = \mathfrak{M}(X)$. The Albanese torus $A(\hat{X})$ is the algebraic reduction of $A(X)$.*

Proof: The family of fibers of an algebraic reduction of a compact complex manifold V is a compact Moishezon subvariety $A_0(V)$ of the Barlet cycle space $\mathcal{C}(V)$ and $A_0(V)$ is invariant under the action of $\text{Aut}^0 V$ on $\mathcal{C}(V)$ ([4]). Therefore after a suitable $\text{Aut}^0 X$ -equivariant desingularization of $A_0(X)$ we get an almost homogeneous Moishezon manifold Z with $\mathfrak{M}(Z) = \mathfrak{M}(X)$, and $A(Z)$ is the algebraic reduction of $A(X)$. Now we study the Albanese bundle $\beta : Z \rightarrow A(Z)$ with typical fiber F and show that $b_1(F) = 0$. Since F is almost homogeneous, the Albanese map $\beta_F : F \rightarrow A(F)$ is a holomorphic fiber bundle. The group $J := \ker \beta_*$ operates on F by restriction $\rho : J \rightarrow \text{Aut } F$ and on $A(F)$ via $\rho_F := (\beta_F)_* \circ \rho$. Note that $Y := \text{Aut}^0 Z \times_{\rho_F} A(F)$ is a homogeneous complex manifold and a holomorphic fiber bundle over $A(Z)$ with fiber $A(F)$. Moreover Z is a holomorphic fiber bundle over Y , induced by β_F ; in particular Y is a Moishezon manifold. A compact homogeneous Moishezon manifold is projective algebraic and is a product of an abelian variety with a projective rational manifold [7]. Therefore Y is an abelian variety and the universality of $A(X)$ yields $Y = A(Z)$, hence $A(F) = \{0\}$. But $\dim_{\mathbb{C}} A(F) = \frac{1}{2}b_1(F)$ for the Moishezon manifold F .

Let ξ be a holomorphic line bundle on Z such that the linear system $H^0(Z, \mathcal{O}(\xi))$ defines a bimeromorphic map of Z onto a projective algebraic variety. Since $\text{Aut}^0 Z$ stabilizes the Chern class $c(\xi)$, the group J stabilizes $c(\xi|_F)$. It follows that $\zeta|_F$ is J -stable, because $H^1(F, \mathcal{O}) = 0$. We may assume that $H^0(F, \mathcal{O}(\zeta|_F))$ defines a bimeromorphic map φ of F onto a closed subvariety $W \subset \mathbb{P}^N$, $N + 1 = \dim_{\mathbb{C}} H^0(F, \mathcal{O}(\zeta|_F))$, such that J -operates as a group of collineations on \mathbb{P}^N , stabilizing W .

The representation $\tilde{\rho} : J \rightarrow \text{Aut } W$, $\tilde{\rho}(h) := \varphi \circ h \circ \varphi^{-1}$, defines a projective algebraic variety $\text{Aut}^0 X \times_{\tilde{\rho}} W$, which is a holomorphic fiber bundle over $A(Z)$, almost homogeneous and bimeromorphically equivalent to Z . After a suitable desingularization we get the desired manifold \tilde{X} . ■

References

- [1] Barth, W. : *Oeljeklaus, E. Über die Albanese-Abbildung einer fasthomogenen Kähler Mannigfaltigkeit.* Math. Ann. **211**, 47–62 (1969)
- [2] Blanchard, A. : *Sur les variétés analytiques complexes.* Ann. Sci. École Norm. Sup. **73**, 157–202 (1958)
- [3] Borel, A. and Remmert, R. : *Über kompakte homogene Kählermannigfaltigkeiten.* Math. Ann. **145**, 429–439 (1962)
- [4] Campana, F. : *Application de l'espace des cycles à la classification biméromorphe des espaces analytiques Kähleriens compacts.* Inst. Élie Cartan, Equipe associée d'Analyse Globale, No. 839 (1980)
- [5] Fischer, G. and Forster, O. : *Ein Endlichkeitssatz für Hyperflächen auf kompakten komplexen Räumen.* J. Reine Angew. Math. **306**, 88–93 (1979)
- [6] Fujiki, A. : *On automorphism groups of compact Kähler manifolds.* Invent. math. **44**, 225–258 (1978)

- [7] Grauert, H. and Remmert, R. : *Über kompakte homogene komplexe Mannigfaltigkeiten*. Arch. Math. **13**, 498–507 (1962)
- [8] Huckleberry, A. and Oeljeklaus, E. : *Homogeneous spaces from a complex analytic viewpoint*. In: Hano, J., et al. (eds.): *Manifolds and Lie Groups. Papers in Honor of Y. Matsushima*. Progress in Mathematics **14**, pp. 159–186. Boston 1981
- [9] Huckleberry, A. and Oeljeklaus, E. : *Classification theorems for almost homogeneous spaces*. Revue de l'institut Élie Cartan **9**, 1–192, (1984)
- [10] Humphreys, J.E. : *Linear Algebraic Groups*. New York 1975
- [11] Kodaira, K. : *On Kähler varieties of restricted type (an intrinsic characterization of algebraic varieties)*. Ann. of Math. **60**, 28–48 (1954)
- [12] Kraft, H. : *Geometrische Methoden in der Invariantentheorie*. Braunschweig 1984
- [13] Lichnerowicz, A. : *Variétés kählériennes à première classe de Chern positive ou nulle*. C.R. Acad. Sc. Paris **268**, 876–880 (1969)
- [14] Lieberman, D. : *Compactness of the Chow scheme, Applications to automorphisms and deformations of Kähler manifolds*. Séminaire F. Norguet (1975–1977), LNM **670**, 140–186 (1978)
- [15] Matsumura, H. : *On algebraic groups of birational transformations*. Rend. Accad. Naz. Lincei, Serie VII, **34**, 151–155 (1963)
- [16] Matsumura, H. and Oort, F. : *Representability of group functors and automorphisms of algebraic schemes*. Invent. Math. **4**, 1–25 (1967)
- [17] Matsusaka, T. : *Polarized varieties, fields of moduli and generalized Kummer varieties of polarized varieties*. Amer. J. Math. **80**, 45–82 (1958)
- [18] Matsushima, Y. : *Sur les espaces homogènes kählériens d'un groupe de Lie réductif*. Nagoya Math. J. **11**, 53–60 (1957)
- [19] Matsushima, Y. : *On Hodge manifolds with zero first Chern class*. J. Differential Geom. **3**, 477–480 (1969)
- [20] Matsushima, Y. : *Holomorphic vector fields on compact Kähler manifolds*. Conf. Board Math. Sci, Reg. Conf. Ser. in Math. **7**, AMS (1971)
- [21] Oeljeklaus, E. : *Fasthomogene Kählermannigfaltigkeit mit verschwindender erster Bettizahl*. Manusk. Math. **7**, 175–183 (1972)
- [22] Rosenlicht, M. : *Some basic theorems on algebraic groups*. Amer. J. Math. **78**, 401–443 (1956)
- [23] Serre, J.-P. : *Géométrie algébrique et géométrie analytique*. Ann. Inst. Fourier **6**, 1–42 (1956)

- [24] Triantaphyllou, D. : *Invariants of finite groups acting on rational function fields.* J. Pure Appl. Algebra **18**, 315–331 (1980)
- [25] Vinberg, E.B. : *Rationality of the field of invariants of a triangular group.* Moscow Univ. Math. Bull. **37**, 2, 27–28 (1982)
- [26] Weil, A. : *On algebraic groups of transformations.* Amer. J. Math. **77**, 355–391 (1955)

received: July 7, 1993

Author:

Prof. Dr. E. Oeljeklaus
Universität Bremen
Fachbereich Mathematik und Informatik
Postfach 33 04 40
28334 Bremen
Germany

WOLFGANG BANNUSCHER; GUNTER TIEDT

On a theorem of Deaconescu¹

ABSTRACT. Consider finite groups having all nontrivial elements of prime order as a subclass of CN-groups in which the centralizers of all nontrivial elements are nilpotent. Then we give in generalization of a Theorem of DEACONESCU [2] necessary and sufficient conditions to be such a group.

1 Introduction

Let CP_1 , CP and CN be the class of finite groups in which the centralizers of all nontrivial elements contain only elements of prime order, of prime power order and are nilpotent, respectively.

Clearly $CP_1 \subset CP \subset CN$, and CP_1, CP consists of exactly those groups having all nontrivial elements of prime order and prime power order, respectively.

DEACONESCU [2] had shown

Theorem 1 *Let G be a CP_1 -group. Then one of the following cases occurs:*

- I. G is a p -group of exponent p .
- II. (a) $|G| = p^a q$, $3 \leq p < q, a \geq 3$, $|F(G)| = p^{a-1}$, $|G : G'| = p$.
 (b) $|G| = p^a q$, $3 \leq p, q, a \geq 1$, $|F(G)| = |G'| = p^a$.
 (c) $|G| = 2^a p$, $p \geq 3$, $a \geq 2$, $|F(G)| = |G'| = 2^a$.
 (d) $|G| = 2p^a$, $p \geq 3$, $a \geq 1$, $|F(G)| = |G'| = p^a$ and $F(G)$ is elementary abelian.
- III. $G \cong A_5$.

Remark: In the main theorem of [2] (IIb) the condition $q < p$ is false because 2.8 of [2] is only true if H is a normal subgroup.

¹Ausgearbeitete Fassung eines Vortrags, gehalten auf dem 3. Bremen - Rostocker Kolloquium, Rostock, 22. und 23. Mai 1992

DEACONESCU asks if the conditions in Theorem 1 (II) are also sufficient. In this note we show that this is not the case in generally. Furthermore we give a new characterization of CP_1 -groups.

2 Result and Examples

In [3] and [1] we can find a classification of all solvable CN -groups and nonsolvable CP -groups. From this we can easily deduce:

Theorem 2 *A group G is a CP -group iff one of the following holds:*

- (i) G is isomorphic with $PSL(2, q)$ with $q = 4, 7, 8, 9, 17$; $PSL(3, 4)$, $Sz(8)$, $Sz(32)$ or M_{10} .
- (ii) G has a nontrivial normal 2-subgroup P and G/P is isomorphic with $PSL(2, 4)$, $PSL(2, 8)$, $Sz(8)$ or $Sz(32)$. Moreover, P is elementary abelian and isomorphic with a direct sum of natural modules.
- (iii) G is a p -group.
- (iv) G is a FROBENIUS group whose kernel is a p -group and the complement is either a cyclic q -group ($q \neq p$) or a generalized quaternion group.
- (v) G is a 3-step group of order $p^a q^b$ (p, q primes, $q > 2$) i.e. $G = O_{pp'}(G)$ and $G \supset O_{pp'}(G)$ with
 - a) $O_{pp'}(G)$ is a FROBENIUS group with kernel $O_p(G)$ and cyclic complement.
 - b) $G/O_p(G)$ is a FROBENIUS group with kernel $O_{pp'}(G)/O_p(G)$.

Now we prove

Theorem 3 *A group G is a CP_1 -group if G is isomorphic with a group of the following list:*

- (i) A_5 ,
- (ii) a p -group of exponent p ,
- (iii) a group of type (iv) in Theorem 2 having only Sylow subgroups of exponent p and q .

Proof. By Theorems 1 and 2 we have only to show that there is no group of type II a) in Theorem 1 in which the Sylow p -subgroup is of exponent p . Assume there is a minimal counterexample G with a Sylow p -subgroup of exponent p and $|G| = p^a q$, $|F(G)| = p^{a-1}$, $|G'| = p^{a-1} q$. Hence G has p -length 2. This is a contradiction to [5] 1.13 Theorem p. 413.

Now we search if the statements of Theorem 1, II are sufficient to be a CP_1 -group. Clearly it is necessary that all Sylow p - and 2-groups must have exponent p or 2, respectively. We cannot omit this supposition:

Example 1. Let G be an abnilpotent group of index system $(3, 13, 3^3)$ (see [7] Theorem 4.2), that means $G = \langle a \rangle \times \langle b \rangle \times GF(3^3)^+$ with $a^3 = b^{13} = 1$, $a^{-1}ba = b^3$, $a^{-1}ca = c^3$, $b^{-1}cb = cw$, where w is a primitive 13th root of unity in $GF(3^3)$. Then we have $|F(G)| = |GF(3^3)^+| = 3^3$ and $|G : G'| = 3$, but a Sylow 3-group of G is not of exponent 3.

Now we add the above supposition in Theorem 1, II. Then statement (a) cannot occur (see Theorem 3). The following example shows that statement (b) is not sufficient for CP_1 -groups.

Example 2. Let $p > 2$ be a prime. Take a natural number n such that there exists a prime $q > 2$ with $qp^n - 1$. Then by [4] (9.24 Satz a), S. 228) there is an extra special p -group P with:

a) $\exp P = p$,

b) $|P| = p^{2n+1}$,

c) $\exists \alpha \in \text{Aut } P$: (i) α acts trivially on $Z(P)$
 (ii) α acts fixed-point-freely on $P/Z(P)$.

Let G be the semidirect product $\langle \alpha \rangle \times P$. It is easy to check that the assertion of II(b) in Theorem 1 is valid. But G has elements of order pq namely $\alpha^i z$ ($i = 1, \dots, q-1, z \in Z(P)$).

But if we know that the Sylow p -subgroup P of G in statement II(b) of Theorem 1 is abelain then the condition is also sufficient: By [6] (Satz 7.13) we have then $G' = P = O_P(G) = [P, G] \times C_P(G) = G' \times C_P(G)$ from which $C_P(G) = 1$ follows and therefore $C_G(x) = \langle x \rangle$ for all q -elements x .

Consequently, the conditions II(c), (d) of Theorem 1 with the above supposition are sufficient to be a CP_1 -group.

References

- [1] Brandl, R. : *Finite groups all of whose elements are of prime power order*. Boll. Un. Mat. Ital. A(5) **18**, No. 3, 491-493 (1981)
- [2] Deanonescu, M. : *Classification of finite groups with all elements of prime order*. Proc. Amer. Math. Soc. **106**, 625-629 (1989)
- [3] Gorenstein, D. : *Finite Groups*. New York 1968
- [4] Huppert, B. : *Endliche Gruppen I*. Berlin 1967
- [5] Huppert, B., and Blackburn, N. : *Finite Groups II*. Berlin 1982
- [6] Kurzweil, H. : *Endliche Gruppen*. Berlin 1977
- [7] Pazderski, G. : *Solvable abnilpotent groups*. Math. Nachr. **117**, 305-321 (1984)

received: September 9, 1993

Authors:

Dr. W. Bannuscher; Dr. G. Tiedt
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

W. B. VASANTHA KANDASAMY

Supermodular subgroups in finite groups

The concept of a modular subgroup originally comes from the theory of lattices. A subgroup M of a group G is called modular in G if the following equalities are satisfied:

$$\begin{aligned}(W \cup M) \cap V &= W \cup (M \cap V) \text{ for all } W, V \text{ in } G \text{ with } W \leq V \text{ and} \\ (W \cup M) \cap V &= M \cup (W \cap V) \text{ for all } W, V \text{ in } G \text{ with } M \leq V\end{aligned}$$

where $X \cup Y = \langle X, Y \rangle$ denotes the subgroup generated by X and Y . Thus M is a modular element of the subgroup lattice in G . Detailed analysis of modular subgroups has been carried out by R. SCHMIDT (cf. [2], [3]).

In this paper we define what are called supermodular subgroups and study them, where the author in [4] and [1] has studied lattices satisfying supermodular identity. It has been proved in [4] and [1] that supermodular lattices form an equational class of lattices lying properly between the equational class of distributive lattices and the equational class of modular lattices.

Definition 1 *Let G be a finite group. A subgroup M of the group G is called supermodular if for all subgroups B, C, D in G we have the following identity to be true:*

$$\begin{aligned}(M \cup B) \cap (M \cup C) \cap (M \cup D) \\ = M \cup \{[(B \cap C) \cap (M \cup D)] \cup [(C \cap D) \cap (M \cup B)] \cup [(B \cap D) \cap (M \cup C)]\} \quad (1)\end{aligned}$$

This identity is known as the supermodular identity.

Proposition 2 *Every supermodular subgroup M of a group G is a modular subgroup of G ; but a modular subgroup M of G need not necessarily be a supermodular subgroup of G .*

Proof. Putting $C = D$ in the supermodular identity we get

$$\begin{aligned}(M \cup B) \cap (M \cup C) &= M \cup \{(B \cap C) \cup [C \cap (M \cup B)] \cup (B \cap C)\} \\ &= M \cup [C \cap (M \cup B)]\end{aligned}$$

to be true for all subgroups B, C in G , which is easily recognized as the condition for a subgroup M to be modular.

Now we prove that every modular subgroup M of G need not necessarily be a supermodular subgroup of G by an example. Let us consider the commutative group $G = \langle a, b, c, d / a^2 = b^2 = c^2 = d^2 = 1 \rangle$. Take $M = \langle a, b, c \rangle$, $B = \langle b, c, d \rangle$, $C = \langle abd \rangle$ and $D = \langle acd \rangle$. The supermodular identity of M would claim

$$(M \cup B) \cap (M \cup C) \cap (M \cup D) \\ = M \cup \{[(B \cap C) \cap (M \cup D)] \cup [(C \cap D) \cap (M \cup B)] \cup [(B \cap D) \cap (M \cup C)]\}.$$

For the L.H.S. we obtain $(M \cup B) \cap (M \cup C) \cap (M \cup D) = G \cap G \cap G = G$, and for the R.H.S. we get $M \cup \{[\{1\} \cap (M \cup D)] \cup [\{1\} \cap (M \cup B)] \cup [\{1\} \cap (M \cup C)]\} = M$. Because obviously L.H.S. \neq R.H.S. we are done.

Theorem 3 *Let G be a finite group in which any two subgroups are comparable. Then every subgroup of G is supermodular.*

Proof. It is known that a finite group in which any two subgroups are comparable necessarily is cyclic of prime power order. Then the lattice of subgroups is a chain, hence distributive and so is supermodular.

Theorem 4 *Let G be a group which has no supermodular subgroups then G has a set of four subgroups M, B_1, C_1, D_1 such that*

$$(M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1) \not\supseteq M$$

while $M > (B_1 \cap C_1) \cap (M \cup D_1)$, $(C_1 \cap D_1) \cap (M \cup B_1)$, $(B_1 \cap D_1) \cap (M \cup C_1)$ holds.

Remark. B_1, C_1, D_1 being distinct subgroups of G , $B_1 \neq C_1$, otherwise $B_1 = B_1 \cap C_1$ and $M \cup B_1 = M \cup (B_1 \cap C_1)$ a contradiction as it will imply equality of $(M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1) = M$.

Proof. As G has no supermodular subgroups there exists subgroups P, X, Y, Z such that

$$(P \cup X) \cap (P \cup Y) \cap (P \cup Z) \not\supseteq P \cup \{[(X \cap Y) \cap (P \cup Z)] \\ \cup [(Y \cap Z) \cap (P \cup X)] \cup [(Z \cap X) \cap (P \cup Y)]\}.$$

Put $M = P \cup \{[(X \cap Y) \cap (P \cup Z)] \cup [(Y \cap Z) \cap (P \cup X)] \cup [(Z \cap X) \cap (P \cup Y)]\}$. Let X, Y and Z denote the subgroups B_1, C_1 and D_1 , respectively, and the subgroup M given by the above relation: $B_1 = X, C_1 = Y, D_1 = Z$.

$$M \cup B_1 = P \cup \{[(X \cap Y) \cap (P \cup Z)] \cup [(Y \cap Z) \cap (P \cup X)] \cup [(Z \cap X) \\ \cap (P \cup Y)]\} \cup X = (P \cup X) \cup (Y \cap Z) \cap (P \cup X) = P \cup X.$$

Similary $M \cup C_1 = P \cup Y$ and $M \cup D_1 = P \cup Z$.

So $(M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1) \not\cong M$. Hence the result.

Theorem 5 *Let G be a group in which every subgroup is modular but having no subgroup which is supermodular. Then G has 4 distinct subgroups M, B, C, D such that*

$$M \cup B = M \cup C = M \cup D \not\cong M,$$

$$M > B \cap C, \quad C \cap D, \quad D \cap B$$

holds.

Proof. By our Theorem 4 we can assert the existence of a set of four elements M, B_1, C_1, D_1 in G such that

$$(M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1) \not\cong M$$

and

$$M > (B_1 \cap C_1) \cap (M \cup D_1), \quad (C_1 \cap D_1) \cap (M \cup B_1), \quad (B_1 \cap D_1) \cap (M \cup C_1).$$

Put

$$B = B_1 \cap (M \cup C_1) \cap (M \cup D_1)$$

$$C = C_1 \cap (M \cup B_1) \cap (M \cup D_1)$$

$$D = D_1 \cap (M \cup B_1) \cap (M \cup C_1).$$

Now consider the four subgroups M, B, C, D ; we have, using the value of B :

$$M \cup B = M \cup B_1 \cap (M \cup C_1) \cap (M \cup D_1) = (M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1)$$

as $(M \cup C_1) \cap (M \cup D_1) > M$ and as every subgroup is modular. Similary

$$M \cup C = (M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1)$$

and

$$M \cup D = (M \cup B_1) \cap (M \cup C_1) \cap (M \cup D_1).$$

So $M \cup B = M \cup C = M \cup D > M$. Hence the result. Now consider $B \cap C$ where B and C are given by the above equations.

$$\begin{aligned} B \cap C &= [B_1 \cap (M \cup C_1) \cap (M \cup D_1)] \cap [C_1 \cap (M \cup B_1) \cap (M \cup D_1)] \\ &= (B_1 \cap C_1) \cap (M \cup D_1) \end{aligned}$$

We get $M > B \cap C$. Similary $M > C \cap D$ and $M > D \cap B$. Hence the theorem.

References

- [1] Iqbalunnisa and Vasantha, W. B. : *Supermodular lattices* J. Madras Univ., Sec. B., **44**, 58-80 (1981)
- [2] Schmidt, R. : *Modulare Untergruppen endlicher Gruppen*. Illinois J. Math. **13**, 358-377 (1969)
- [3] Schmidt, R. : *Modular subgroups of finite groups II*. Illinois J. Math. **14**, 344-362 (1969)
- [4] Vasantha, W. B. : *Studies in algebra with special reference to simple modular lattices and supermodular lattices*. Ph. D. Thesis, University of Madras, 1980

received: November 1, 1991

revised: November 11, 1993

Author:

Dr. W. B. Vasantha Kandasamy
Department of Mathematics
Indian Institute of Technology
Madras - 600 036
India

FRIEDRICH LIESE; ANDREAS WIENKE

Asymptotische Normalität von Maximum – Likelihood – Schätzern in inhomogenen Wienerprozessen¹

1 Einleitung

Ein stochastischer Prozess $W(t)$ ($0 \leq t \leq T$) heißt inhomogener Wienerprozeß, falls er im Nullpunkt startet ($W(0) = 0$), seine Realisierungen stetige Funktionen sind und er unabhängige und normalverteilte Zuwächse besitzt. Solche Prozesse beschreiben z.B. die Abnutzung technischer Systeme. Man geht hierbei von der Vorstellung aus, daß die Abnutzung $W(t) - W(s)$ im Zeitraum $[s, t]$ unabhängig von der Vergangenheit ist und die Abnutzung zum Zeitpunkt t sich als Summe aus der Abnutzung $W(s)$ zum Zeitpunkt s und der in der Zeit von s bis t erfolgten Abnutzung darstellt. Das wahrscheinlichkeitstheoretische Verhalten eines inhomogenen Wienerprozesses $W(t)$ wird vollständig durch die Erwartungswertfunktion $m(t) = EW(t)$ und die Varianzfunktion $b(t) = V(W(t))$ beschrieben. Bei der Modellierung des Abnutzungsverhaltens von Systemen geht man jetzt davon aus, daß die Funktionen $m(t, \vartheta)$ und $b(t, \vartheta)$ bis auf einen endlichdimensionalen Parameter $\vartheta \in \Theta \subset \mathbb{R}^k$ bekannt sind, der geschätzt werden soll. Sei P_ϑ ($\vartheta \in \Theta$) die Verteilung von W im Raum der stetigen Funktionen auf $[0, T]$. Hängt die Varianzfunktion $b(t, \vartheta)$ von ϑ ab, so sind für verschiedene Parameter ϑ_1, ϑ_2 die Verteilungen P_{ϑ_1} und P_{ϑ_2} gegenseitig singulär (siehe Jacod, Shiryaev [2], Kap. IV §4c). Eine sinnvolle Likelihoodtheorie ist also nicht möglich. Man vergrößert deshalb das Modell und betrachtet nur Beobachtungen an diskreten Zeitpunkten $0 = t_{0,n} < t_{1,n} < \dots < t_{k_n,n} = T$, wobei die Anzahl k_n gegen Unendlich strebt. Die Voraussetzung, daß nur Beobachtungen an diskreten Zeitpunkten vorliegen, ist auch vom praktischen Standpunkt realistischer. Zusätzlich hat man in vielen Fällen wiederholte unabhängige Beobachtungen vorliegen. Wir gehen deshalb davon aus, daß m_n unabhängig identisch verteilte Wienerprozesse W_1, \dots, W_{m_n} zu den Zeitpunkten $t_{i,n}$ beobachtet werden.

¹Ausgearbeitete und erweiterte Fassung eines Vortrages, gehalten auf dem Bremen – Rostocker Statistik – Seminar, Rostock, 5. - 7. März 1992

Aus diesen $k_n m_n$ Daten soll $\vartheta \in \Theta$ geschätzt werden. Allgemeine Kriterien für die Konsistenz von Maximum – Likelihood – Schätzern (MLS) $\widehat{\vartheta}_n$ in diesem Modell einschließlich Geschwindigkeitsabschätzungen wurden in Liese, Wienke [3] abgeleitet.

In der vorliegenden Arbeit wird die Verteilung von $\sqrt{N_n}(\widehat{\vartheta}_n - \vartheta_0)$ für $n \rightarrow \infty$ untersucht, wobei $N_n \rightarrow \infty$ gilt. Hierbei bezeichnet ϑ_0 den unbekanntem wahren Parametervektor. Ist die Varianzfunktion b von ϑ unabhängig und hängt nur die Erwartungswertfunktion m von ϑ ab, so ist $N_n = m_n$ zu wählen und in Theorem 2 wird gezeigt, daß unter gewissen schwachen Regularitätsvoraussetzungen an b, m und die Beobachtungszeitpunkte $t_{i,n}$ die Verteilung von $\sqrt{m_n}(\widehat{\vartheta}_n - \vartheta_0)$ schwach gegen eine k -dimensionale Normalverteilung mit dem Erwartungswertvektor Null strebt, wobei in die Kovarianzmatrix dieser Normalverteilung Ableitungen der Funktion $m(t, \vartheta)$ nach den Komponenten von ϑ eingehen. Von der Folge der Beobachtungszeitpunkte $t_{i,n}$ wird hierbei nur vorausgesetzt, daß $\text{Max}_{1 \leq i \leq k_n} |t_{i,n} - t_{i-1,n}| \xrightarrow{n \rightarrow \infty} 0$ gilt. Die auftretende Kovarianzmatrix hängt auch nicht von der speziell gewählten Folge $t_{i,n}$ ab.

Qualitativ anders ist die Situation, wenn die Varianzfunktion b von ϑ abhängt. Dann gilt $N_n = k_n m_n$. Man hat also eine wesentlich größere Konvergenzgeschwindigkeit, und es ist hier auch möglich, daß die Anzahl der wiederholten Beobachtungen nicht gegen Unendlich strebt. Es ist sogar $m_n = 1$ zugelassen, wobei die Schätzungen auf einer wachsenden Zahl von Beobachtungen einer Realisierung an diskreten Zeitpunkten basieren. Theorem 1 besagt, daß die Verteilung von $\sqrt{k_n m_n}(\widehat{\vartheta}_n - \vartheta_0)$ gegen eine k -dimensionale Normalverteilung mit dem Erwartungswertvektor Null konvergiert. In die zugehörige Kovarianzmatrix gehen Ableitungen von $b(t, \vartheta)$ nach den Komponenten von ϑ ein. Die Erwartungswertfunktion m geht nicht in die Kovarianzmatrix ein.

Im Gegensatz zur Situation in Theorem 2 geht die spezielle Form der Beobachtungszeitpunkte $t_{i,n}$ ($1 \leq i \leq k_n$, $n = 1, 2, \dots$) ein. Sei $\mu_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \delta_{t_{i,n}}$ das zum Versuchsplan $\{t_{i,n}\}$ gehörige empirische Maß und es gelte $\mu_n \Rightarrow \mu$ im Sinne der schwachen Konvergenz. In die Kovarianzmatrix der Grenzverteilung in Theorem 1 geht dann das Maß μ ein.

2 Modellvoraussetzungen und Resultate

Ein stochastischer Prozeß $W(t)$ ($0 \leq t \leq T$) heißt inhomogener Wienerprozeß mit der Erwartungswertfunktion $m(t)$ und der Varianzfunktion $b(t)$, falls W ein stetiger Gaußprozeß mit unabhängigen Zuwächsen ist, so daß $W(0) = 0$ gilt und die Zuwächse $W(t) - W(s)$ für $0 \leq s < t \leq T$ eine Normalverteilung mit dem Erwartungswert $m(t) - m(s)$ und der Varianz $b(t) - b(s)$ besitzen. m ist dann notwendigerweise eine stetige Funktion mit $m(0) = 0$ und b eine stetige nichtfallende Funktion mit $b(0) = 0$. Wir nehmen jetzt an, daß die Erwartungswertfunktion und die Varianzfunktion nur bis auf einen endlichdimensionalen

Parameter $\vartheta \in \Theta \subseteq \mathbb{R}^k$ bekannt sind. ϑ soll durch Messungen zu den diskreten Zeitpunkten $0 = t_{0,n} < t_{1,n} < \dots < t_{k_n,n} = T$ geschätzt werden.

In vielen Fällen liegen zusätzlich unabhängige Beobachtungen W_1, \dots, W_{m_n} des Prozesses W vor. Wegen $W_j(0) = 0$ können wir statt mit den Werten $W_j(t_{i,n})$ gleichberechtigt mit den Differenzen $W_j(t_{i,n}) - W_j(t_{i-1,n})$ arbeiten, um den Parameter ϑ zu schätzen. Der Übergang zu den stochastisch leichter handhabbaren Differenzen ergibt keinen Informationsverlust.

Sei $N(\mu, \sigma^2)$ die Normalverteilung auf den Borelmengen des \mathbb{R}^1 mit dem Erwartungswert μ und der Varianz σ^2 . \mathcal{B}^m ist die σ -Algebra der Borelmengen des \mathbb{R}^m . Wir bezeichnen mit P_ϑ^n das folgende Produktmaß:

$$P_\vartheta^n = \prod_{j=1}^{m_n} \prod_{i=1}^{k_n} N(m(t_{i,n}, \vartheta) - m(t_{i-1,n}, \vartheta), b(t_{i,n}, \vartheta) - b(t_{i-1,n}, \vartheta)). \quad (1)$$

Dabei kennzeichnet P_ϑ^n die Verteilung der m_n identisch verteilten Vektoren der Zuwächse $W_j(t_{1,n}), W_j(t_{2,n}) - W_j(t_{1,n}), \dots, W_j(t_{k_n,n}) - W_j(t_{k_n-1,n})$, die jeweils die k_n dimensionale Normalverteilung

$$\prod_{i=1}^{k_n} N(m(t_{i,n}, \vartheta) - m(t_{i-1,n}, \vartheta), b(t_{i,n}, \vartheta) - b(t_{i-1,n}, \vartheta))$$

als Verteilung besitzen. Insgesamt wird das geschilderte Schätzproblem beschrieben als Schätzproblem für die Folge der statistischen Experimente E_n mit

$$E_n = (\mathbb{R}^{k_n m_n}, \mathcal{B}^{k_n m_n}, P_\vartheta^n, \vartheta \in \Theta). \quad (2)$$

Die Meßzeitpunkte $t_{i,n}$ sollen schwachen Regularitätsvoraussetzungen genügen. Es sei

$$(V1) \quad t_{0,n} = 0, \quad t_{k_n,n} = T, \quad \max_{1 \leq i \leq k_n} (t_{i,n} - t_{i-1,n}) \xrightarrow{n \rightarrow \infty} 0.$$

Um das asymptotische Verhalten der Folge der Versuchspläne $\{t_{i,n}, 0 \leq i \leq k_n\}$ beschreiben zu können, ordnen wir ihnen empirische Maße zu. Es sei

$$\mu_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \delta_{t_{i,n}},$$

wobei δ_x das δ -Maß im Punkt x bezeichnet. Wir setzen voraus, daß

$$(V2) \quad \mu_n \implies \mu, \quad n \rightarrow \infty,$$

gilt, wobei \implies das Symbol für die schwache Konvergenz von Verteilungen auf $[0, T]$ bezeichnet. Von der Varianzfunktion $b(t, \vartheta)$ setzen wir voraus, daß sie folgender Bedingung genügt:

$$(V3) \quad t \longmapsto b(t, \vartheta) \text{ ist stetig differenzierbar in } [0, T] \text{ für alle } \vartheta \in \Theta \text{ und es gilt } \dot{b}(t, \vartheta) = \frac{\partial}{\partial t} b(t, \vartheta) > 0 \text{ für alle } 0 \leq t \leq T \text{ und alle } \vartheta \in \Theta.$$

Die Voraussetzung $\dot{b}(t, \vartheta) > 0$ erzwingt insbesondere, daß die Differenzen $W_j(t_{i,n}) - W_j(t_{i-1,n})$ stets positive Varianzen besitzen und die Normalverteilung

$$N(m(t_{i,n}, \vartheta) - m(t_{i-1,n}, \vartheta), b(t_{i,n}, \vartheta) - b(t_{i-1,n}, \vartheta))$$

nicht zu einer δ -Verteilung entartet. Insbesondere sind deshalb die P_ϑ^n aus (1) durch das $k_n m_n$ -dimensionale Lebesguesche Maß λ dominiert. Für eine Funktion $f(t, \vartheta)$ werden durch $\dot{f}(t, \vartheta)$ die Ableitung nach t und durch $f_{u,v}, f_{u,v}, \dots$ ($u, v = 1, \dots, k$) die Ableitungen nach den Komponenten des Vektors $(\vartheta_1, \dots, \vartheta_k)$ bezeichnet. Insbesondere ist z.B.

$$\dot{f}_{u,v} = \frac{\partial^3}{\partial t \partial \vartheta_u \partial \vartheta_v} f.$$

Θ^0 sei das Innere von Θ und Θ^c der Abschluß der Menge Θ . Ist f eine auf $[0, T] \times \Theta$ definierte reelle Funktion, so sagen wir, daß sie der Bedingung (V4) genügt, falls gilt:

$$(V4) \quad \begin{aligned} t \mapsto f(t, \vartheta) \text{ ist differenzierbar nach } t \forall \vartheta \in \Theta. \dot{f} \text{ ist in } \Theta^0 \text{ zweimal} \\ \text{stetig differenzierbar nach allen Komponenten von } \vartheta = (\vartheta_1, \dots, \vartheta_k) \\ \text{und } (t, \vartheta) \mapsto \dot{f}_{u,v}(t, \vartheta) \text{ ist stetig in } [0, T] \times \Theta^0 \text{ für alle } 1 \leq u, v \leq k. \end{aligned}$$

$\Sigma(\vartheta)$ sei die $k \times k$ Matrix mit den Elementen

$$\sigma_{u,v}(\vartheta) = \frac{1}{2} \int_{[0, T]} \frac{\dot{b}_u(t, \vartheta) \dot{b}_v(t, \vartheta)}{(\dot{b}(t, \vartheta))^2} \mu(dt).$$

Eine Folge meßbarer Abbildungen $\widehat{\vartheta}_n : \mathbf{R}^{k_n m_n} \rightarrow \Theta^c$ heißt konsistente Schätzfolge, falls

$$P_\vartheta^n \left(\|\widehat{\vartheta}_n - \vartheta\| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \quad (3)$$

für alle $\varepsilon > 0$ gilt. Ist X ein auf dem Wahrscheinlichkeitsraum E_n definierter zufälliger Vektor, so bezeichnet $\mathcal{L}_\vartheta(X) = P_\vartheta^n \circ X^{-1}$ die Verteilung von X .

Theorem 1 Die Funktionen m und b mögen der Voraussetzung (V4) genügen, b erfülle zusätzlich (V3). Für die Folge der Versuchspläne $\{t_{i,n}, 0 \leq i \leq k_n\}$ gelte (V1) und (V2). $\widehat{\vartheta}_n$ sei eine konsistente Folge von ML -Schätzern bez. der Folge E_n aus (2). ϑ_0 sei innerer Punkt von Θ und $\Sigma(\vartheta_0)$ sei regulär ($\text{Det}(\Sigma(\vartheta_0)) \neq 0$). Dann gilt

$$\mathcal{L}_{\vartheta_0} \left(\sqrt{k_n m_n} (\widehat{\vartheta}_n - \vartheta_0) \right) \implies N \left(0, \Sigma^{-1}(\vartheta_0) \right),$$

wobei $N \left(0, \Sigma^{-1}(\vartheta_0) \right)$ die k -dimensionale Normalverteilung mit dem Erwartungswertvektor 0 und der Kovarianzmatrix $\Sigma^{-1}(\vartheta_0)$ ist.

Theorem 1 läßt die Möglichkeit zu, daß alle m_n gleich 1 sind, also keine wiederholten Beobachtungen vorliegen.

In diesem Falle wird die Folge der Maximum - Likelihood - Schätzer aus einer Realisierung des Prozesses $W(t)$ ($0 \leq t \leq T$) ermittelt, wobei die Beobachtungen an einer wachsenden Zahl von Zeitpunkten im Intervall $[0, T]$ vorgenommen werden. Die Existenz von konsistenten Schätzern ergibt sich aus der Tatsache, daß die Verteilungen P_ϑ , P_η des Prozesses W im Raum der stetigen Funktionen für unterschiedliche Parameterwerte ϑ, η gegenseitig singular sind. Hieraus folgt, daß unter den Voraussetzungen (V1) und (V2) die Verteilungen $\{P_\vartheta^n\}$ und $\{P_\eta^n\}$ vollständig trennbar sind in dem Sinne, daß $A_n \in \mathcal{B}^{k_n m_n}$ existieren mit $P_\vartheta^n(A_n) \xrightarrow{n \rightarrow \infty} 1$, $P_\eta^n(A_n) \xrightarrow{n \rightarrow \infty} 0$, falls $\vartheta \neq \eta$.

Obwohl man konsistente Schätzungen bereits aus einer kontinuierlichen Realisierung erhält, erhöhen zusätzliche unabhängige Wiederholungen des Prozesses die Konvergenzgeschwindigkeit der Folge von Schätzern. Dies drückt sich im Faktor $\sqrt{m_n}$ im obigen Grenzwertsatz aus.

Ist die Varianzfunktion $b(t, \vartheta)$ von ϑ unabhängig, hängt also P_ϑ^n nur über $m(t, \vartheta)$ von ϑ ab, so ist Theorem 1 wegen $\text{Det}(\widehat{\Sigma}(\vartheta_0)) = 0$ nicht anwendbar. Ist $b(t, \vartheta)$ von ϑ unabhängig, so existieren konsistente Schätzer für ϑ nur, falls $m_n \rightarrow \infty$ gilt, weil die zugehörigen Verteilungen P_ϑ , P_η nicht singular sind. Damit lassen sich die Folgen $\{P_\vartheta^n\}$, $\{P_\eta^n\}$ nur dann vollständig trennen, wenn $m_n \rightarrow \infty$ gilt.

Wir bezeichnen durch $\widehat{\Sigma}(\vartheta_0)$ die $k \times k$ Matrix mit den Elementen

$$\widehat{\sigma}_{uv}(\vartheta_0) = \int_0^T \frac{\dot{m}_u(t, \vartheta_0) \dot{m}_v(t, \vartheta_0)}{\dot{b}(t)} dt.$$

Theorem 2 Die Varianzfunktion $b(t)$ des inhomogenen Wienerprozesses W sei unabhängig von ϑ und in $[0, T]$ stetig differenzierbar. Es gelte $\dot{b}(t) > 0$ ($0 \leq t \leq T$). Zusätzlich genüge $\dot{m}_{u,v}(t, \vartheta)$ einer Hölderbedingung bezüglich ϑ in einer Umgebung $U(\vartheta_0)$ von ϑ_0 mit einer von t unabhängigen Hölderkonstanten c für alle u, v :

$$|\dot{m}_{u,v}(t, \vartheta_1) - \dot{m}_{u,v}(t, \vartheta_2)| \leq c \|\vartheta_1 - \vartheta_2\|^\alpha, \quad \alpha > 0, \quad \forall \vartheta_1, \vartheta_2 \in U(\vartheta_0).$$

Es seien die Voraussetzungen (V1) und (V2) erfüllt. ϑ_0 sei ein innerer Punkt von Θ und es gelte $\text{Det}(\widehat{\Sigma}(\vartheta_0)) \neq 0$ sowie $m_n \rightarrow \infty$. Dann gilt für jede konsistente Folge $\widehat{\vartheta}_n$ von Maximum - Likelihood - Schätzern

$$\mathcal{L}_{\vartheta_0}(\sqrt{m_n}(\widehat{\vartheta}_n - \vartheta_0)) \Rightarrow N\left(0, \widehat{\Sigma}^{-1}(\vartheta_0)\right).$$

3 Beweise

Als Vorbereitung auf die eigentlichen Beweise von Theorem 1 und Theorem 2 leiten wir zunächst einige technische Lemmata ab. Zur Abschätzung von Stetigkeitsmodulen benötigen wir ein Ergebnis aus Ibragimov, Hasminskij [1]. Vom dort angegebenen Satz 19 im Anhang 1 brauchen wir lediglich einen Spezialfall, den wir im folgenden Lemma formulieren.

Lemma 1 $K \subseteq \mathbb{R}^k$ sei eine abgeschlossene und beschränkte Menge und $\xi(\vartheta)$ ($\vartheta \in K$) ein zufälliger Prozeß mit stetigen Realisierungen. Es möge Zahlen $m \geq r > k$ und H derart geben, daß

$$E|\xi(\vartheta)|^m \leq H \quad (4)$$

$$E|\xi(\vartheta_1) - \xi(\vartheta_2)|^m \leq \|\vartheta_1 - \vartheta_2\|^r \quad (5)$$

für alle $\vartheta_i \in K, i = 1, 2$, gilt. Dann existiert eine Konstante B_0 (die nur von m, r, k und dem Durchmesser von K abhängt) derart, daß

$$E \sup_{\|\vartheta_1 - \vartheta_2\| \leq h} |\xi(\vartheta_1) - \xi(\vartheta_2)| \leq B_0 H^m h^{\frac{r-k}{m}}.$$

Lemma 2 f und g seien stetige Funktionen in $[0, T]$. Die Folge $\{t_{i,n}, 0 \leq i \leq k_n\}$ genüge den Voraussetzungen (V1) und (V2). Dann gilt für beliebige $s_{i,n}^*, t_{i,n}^* \in [t_{i-1,n}, t_{i,n}]$

$$\frac{1}{k_n} \sum_{i=1}^{k_n} f(s_{i,n}^*) g(t_{i,n}^*) \rightarrow \int_{[0,T]} f(t) g(t) \mu(dt).$$

Beweis: Die Beziehung

$$\frac{1}{k_n} \sum_{i=1}^{k_n} f(t_{i,n}) g(t_{i,n}) = \int_{[0,T]} f(t) g(t) \mu_n(dt) \xrightarrow{n \rightarrow \infty} \int_{[0,T]} f(t) g(t) \mu(dt) \quad (6)$$

folgt aus der schwachen Konvergenz $\mu_n \Rightarrow \mu$. Die gleichmäßige Stetigkeit von f und g auf $[0, T]$ ergibt wegen (V1)

$$\max_{1 \leq i \leq k_n} \left| f(t_{i,n}) g(t_{i,n}) - f(s_{i,n}^*) g(t_{i,n}^*) \right| \xrightarrow{n \rightarrow \infty} 0,$$

woraus wegen (6) die Behauptung folgt. ■

Bei festem $\vartheta = (\vartheta_1, \dots, \vartheta_k) \in \Theta$ setzen wir

$$\Delta m(t_{i,n}, \vartheta) = m(t_{i,n}, \vartheta) - m(t_{i-1,n}, \vartheta)$$

$$\Delta b(t_{i,n}, \vartheta) = b(t_{i,n}, \vartheta) - b(t_{i-1,n}, \vartheta)$$

$$\Delta m_u(t_{i,n}, \vartheta) = m_u(t_{i,n}, \vartheta) - m_u(t_{i-1,n}, \vartheta).$$

Die Größen $\Delta b_u(t_{i,n}, \vartheta)$, $\Delta m_{u,v}(t_{i,n}, \vartheta)$ und $\Delta b_{u,v}(t_{i,n}, \vartheta)$ sind dann sinngemäß definiert.

Lemma 3 Die Funktionen m und b mögen (V4) erfüllen. Für b gelte (V3). Weiterhin seien (V1) und (V2) erfüllt. Dann gilt für $\vartheta_0 \in \Theta^0$

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \frac{\Delta b_u(t_{i,n}, \vartheta_0) \Delta b_v(t_{i,n}, \vartheta_0)}{2(\Delta b(t_{i,n}, \vartheta_0))^2} \xrightarrow{n \rightarrow \infty} \frac{1}{2} \int_{[0, T]} \frac{\dot{b}_u(t, \vartheta_0) \dot{b}_v(t, \vartheta_0)}{\dot{b}^2(t, \vartheta_0)} \mu(dt)$$

und

$$\frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} \xrightarrow{n \rightarrow \infty} 0.$$

Beweis: Wir setzen

$$f(t) = \frac{\dot{b}_u(t, \vartheta_0)}{\dot{b}(t, \vartheta_0)}, \quad g(t) = \frac{\dot{b}_v(t, \vartheta_0)}{\dot{b}(t, \vartheta_0)}.$$

Nach Voraussetzung sind f und g stetige Funktionen in $[0, T]$. Aus dem verallgemeinerten Mittelwertsatz der Differentialrechnung folgt mit gewissen $s_{i,n}^*, t_{i,n}^* \in [t_{i-1,n}, t_{i,n}]$

$$\sum_{i=1}^{k_n} \frac{\Delta b_u(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} \frac{\Delta b_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} = \sum_{i=1}^{k_n} f(s_{i,n}^*) g(t_{i,n}^*).$$

Die Anwendung von Lemma 2 ergibt

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \frac{\Delta b_u(t_{i,n}, \vartheta_0) \Delta b_v(t_{i,n}, \vartheta_0)}{2(\Delta b(t_{i,n}, \vartheta_0))^2} \xrightarrow{n \rightarrow \infty} \frac{1}{2} \int_{[0, T]} \frac{\dot{b}_u(t, \vartheta_0) \dot{b}_v(t, \vartheta_0)}{\dot{b}^2(t, \vartheta_0)} \mu(dt).$$

Es sei

$$c = \sup_{0 \leq t \leq T} |\dot{m}(t, \vartheta_0)| \quad \text{und} \quad d = \sup_{0 \leq t \leq T} \left| \frac{1}{\dot{b}(t, \vartheta_0)} \right|.$$

Nach Voraussetzung gilt $c < \infty$ und $d < \infty$. Damit haben wir

$$\left| \frac{1}{k_n} \sum_{i=1}^{k_n} \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} \right| \leq \frac{c^2}{d} \max_{1 \leq i \leq k_n} |t_{i,n} - t_{i-1,n}| \xrightarrow{n \rightarrow \infty} 0, \quad (7)$$

und dies ergibt die Behauptung. ■

Lemma 4 Die Funktionen b und m mögen die Voraussetzungen in Theorem 2 erfüllen. Weiterhin gelte (V1). Dann folgt mit $\Delta b(t_{i,n}) := b(t_{i,n}) - b(t_{i-1,n})$

$$\sum_{i=1}^{k_n} \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} \rightarrow \int_0^T \frac{\dot{m}_u(t, \vartheta_0) \dot{m}_v(t, \vartheta_0)}{\dot{b}(t)} dt.$$

Beweis: Aus dem verallgemeinerten Mittelwertsatz der Differentialrechnung folgt mit gewissen $s_{i,n}^*$, $t_{i,n}^* \in [t_{i-1,n}, t_{i,n}]$

$$\sum_{i=1}^{k_n} \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} = \sum_{i=1}^{k_n} \frac{\dot{m}_u(s_{i,n}^*, \vartheta_0) \dot{m}_v(t_{i,n}^*, \vartheta_0)}{\dot{b}(s_{i,n}^*)} (t_{i,n} - t_{i-1,n}).$$

Die gleichmäßige Stetigkeit von $t \mapsto \dot{m}_v(t, \vartheta_0)$ ergibt

$$\max_{1 \leq i \leq k_n} \left| \dot{m}_v(s_{i,n}^*, \vartheta_0) - \dot{m}_v(t_{i,n}^*, \vartheta_0) \right| \xrightarrow{n \rightarrow \infty} 0.$$

Damit erhalten wir die Behauptung aus der Riemannintegrierbarkeit von

$$t \mapsto \frac{\dot{m}_u(t, \vartheta_0) \dot{m}_v(t, \vartheta_0)}{\dot{b}(t)}. \quad \blacksquare$$

Lemma 5 $Y_{i,j,n}$ ($1 \leq i \leq k_n$, $1 \leq j \leq m_n$, $n = 1, 2, \dots$) seien unabhängige standardnormalverteilte Zufallsvariable. $c_{i,n}$, $d_{i,n}$ ($1 \leq i \leq k_n$, $n = 1, 2, \dots$) seien reelle Zahlen mit

$$D = \sup_{i,n} (|c_{i,n}| + |d_{i,n}|) < \infty \quad (8)$$

und der Eigenschaft, daß

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{k_n} \sum_{i=1}^{k_n} (2c_{i,n}^2 + d_{i,n}^2) \quad (9)$$

existiert und $\sigma^2 > 0$ erfüllt ist. Dann gilt

$$\mathcal{L} \left(\frac{1}{\sqrt{k_n m_n}} \sum_{j=1}^{m_n} \sum_{i=1}^{k_n} \left((Y_{i,j,n}^2 - 1) c_{i,n} + Y_{i,j,n} d_{i,n} \right) \right) \Rightarrow N(0, \sigma^2).$$

Beweis: Es bezeichne B_n^2 die Varianz von $\sum_{j=1}^{m_n} \sum_{i=1}^{k_n} (Y_{i,j,n}^2 - 1) c_{i,n} + Y_{i,j,n} d_{i,n}$. Eine einfache Rechnung zeigt

$$B_n^2 = m_n \sum_{i=1}^{k_n} (2c_{i,n}^2 + d_{i,n}^2).$$

Aufgrund des zentralen Grenzwertsatzes für das Serienschema ist nur noch die Ljapunowsche Bedingung

$$\lim_{n \rightarrow \infty} \frac{1}{B_n^3} \sum_{j=1}^{m_n} \sum_{i=1}^{k_n} E \left| (Y_{i,j,n}^2 - 1) c_{i,n} + Y_{i,j,n} d_{i,n} \right|^3 = 0 \quad (10)$$

nachzuweisen. Aus $\mathcal{L}(Y_{i,j,n}) = N(0, 1)$ und (8) erhält man

$$\sup_{i,n} E \left| (Y_{i,j,n}^2 - 1) c_{i,n} + Y_{i,j,n} d_{i,n} \right|^3 < \infty.$$

Wegen (9) gilt für alle hinreichend großen n

$$B_n^3 \geq \left(\frac{\sigma^2}{2}\right)^{\frac{3}{2}} (m_n k_n)^{\frac{3}{2}}.$$

Aus diesen beiden Ungleichungen folgt wegen $k_n \rightarrow \infty$ die Bedingung (10) und damit die Behauptung. ■

Lemma 6 Die Zufallsvariablen $Y_{i,j,n}$ seien wie in Lemma 5 gewählt. $S_{i,n}$ ($1 \leq i \leq k_n$, $n = 1, 2, \dots$) seien reelle Zahlen mit $R = \sup_{i,n} |S_{i,n}| < \infty$. Dann gilt

$$E \left(\frac{1}{k_n} \sum_{i=1}^{k_n} (Y_{i,j,n}^2 - 1) S_{i,n} \right)^2 \leq \frac{2R^2}{k_n}.$$

Beweis: Das Lemma folgt aus der Unabhängigkeit der $Y_{i,j,n}$ und $EY_{i,j,n}^2 = 1$, $EY_{i,j,n}^4 = 3$. ■

Wir untersuchen jetzt den Logarithmus der Likelihoodfunktion im Experiment E_n , das in (2) erklärt wurde. Sei $L_n(\vartheta) = \ln \frac{dP_n}{d\lambda}$, wobei λ das $k_n m_n$ -dimensionale Lebesguesche Maß ist. Dann gilt

$$L_n(x_{1,1}, \dots, x_{k_n, m_n}, \vartheta) = \sum_{j=1}^{m_n} L_{j,n}(x_{1,j}, \dots, x_{k_n, j}, \vartheta) \quad (11)$$

mit

$$L_{j,n}(x_{1,j}, \dots, x_{k_n, j}, \vartheta) = -\frac{k_n}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^{k_n} \left[\ln(\Delta b(t_{i,n}, \vartheta)) + \frac{(x_{i,j} - \Delta m(t_{i,n}, \vartheta))^2}{\Delta b(t_{i,n}, \vartheta)} \right]. \quad (12)$$

Wir bezeichnen mit $X_{i,j} : \mathbb{R}^{k_n m_n} \rightarrow \mathbb{R}^1$ ($1 \leq i \leq k_n$, $j = 1 \dots m_n$) die Projektionen auf die Koordinaten. Fassen wir sie als Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\mathbb{R}^{k_n m_n}, \mathcal{B}^{k_n m_n}, P_{\vartheta_0}^n)$ auf, so sind sie unabhängige Zufallsvariable mit

$$\mathcal{L}_{\vartheta_0}(X_{i,j}) = N(\Delta m(t_{i,n}, \vartheta_0), \Delta b(t_{i,n}, \vartheta_0)).$$

Deshalb ist bei festem n die Größe $L_n(\vartheta)$ eine Summe aus den unabhängigen und identisch verteilten Zufallsvariablen $L_{j,n}(\vartheta)$.

Erwartungswerte, Varianzen und Kovarianzen, die sich auf das Maß $P_{\vartheta_0}^n$ beziehen, werden durch E_{ϑ_0} , V_{ϑ_0} , cov_{ϑ_0} bezeichnet. Wir berechnen die Ableitungen der Log - Likelihoodfunktion nach den Komponenten von ϑ . Mit der Bezeichnung

$$L_{n,u}(\vartheta) = \frac{\partial}{\partial \vartheta_u} L_n(\vartheta)$$

gilt

$$L_{n,u}(\vartheta) = \sum_{j=1}^{m_n} L_{j,n,u}(\vartheta) \quad (13)$$

$$L_{n,u,v}(\vartheta) = \sum_{j=1}^{m_n} L_{j,n,u,v}(\vartheta), \quad (14)$$

wobei

$$\begin{aligned} L_{j,n,u}(\vartheta) &= \sum_{i=1}^{k_n} \left[\frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta))^2}{2(\Delta b(t_{i,n}, \vartheta))^2} \Delta b_u(t_{i,n}, \vartheta) \right. \\ &\quad \left. - \frac{\Delta b_u(t_{i,n}, \vartheta)}{2\Delta b(t_{i,n}, \vartheta)} + \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta)}{\Delta b(t_{i,n}, \vartheta)} \Delta m_u(t_{i,n}, \vartheta) \right] \quad (15) \\ &= \sum_{i=1}^{k_n} Y_{i,j,u}(\vartheta) \end{aligned}$$

$$\begin{aligned} L_{j,n,u,v}(\vartheta) &= \sum_{i=1}^{k_n} \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta))^2}{\Delta b(t_{i,n}, \vartheta)} \left[\frac{\Delta b_{u,v}(t_{i,n}, \vartheta)}{2\Delta b(t_{i,n}, \vartheta)} - \frac{\Delta b_u(t_{i,n}, \vartheta) \Delta b_v(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^2} \right] \\ &\quad + \sum_{i=1}^{k_n} \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^{1/2}} \left[\frac{\Delta m_{u,v}(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^{1/2}} - \frac{\Delta m_u(t_{i,n}, \vartheta) \Delta b_v(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^{3/2}} \right. \\ &\quad \left. - \frac{\Delta m_u(t_{i,n}, \vartheta) \Delta b_v(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^{3/2}} \right] \\ &\quad - \sum_{i=1}^{k_n} \left[\frac{\Delta m_u(t_{i,n}, \vartheta) \Delta m_v(t_{i,n}, \vartheta)}{\Delta b(t_{i,n}, \vartheta)} + \frac{\Delta b_{u,v}(t_{i,n}, \vartheta)}{2\Delta b(t_{i,n}, \vartheta)} - \frac{\Delta b_u(t_{i,n}, \vartheta) \Delta b_v(t_{i,n}, \vartheta)}{2(\Delta b(t_{i,n}, \vartheta))^2} \right]. \quad (16) \end{aligned}$$

Wir untersuchen die Momente von $L_{j,n,u}(\vartheta)$. Offensichtlich gilt $E_{\vartheta_0} L_{j,n,u}(\vartheta_0) = 0$. Aus der Unabhängigkeit von X_{i,j_1} und X_{i,j_2} für $j_1 \neq j_2$ und

$$\mathcal{L}_{\vartheta_0} \left(\frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}, \vartheta_0))^{1/2}} \right) = N(0, 1)$$

ergibt sich

$$\text{cov}_{\vartheta_0}(L_{j,n,u}(\vartheta_0), L_{j,n,v}(\vartheta_0)) = \sum_{i=1}^{k_n} \text{cov}_{\vartheta_0}(Y_{i,j,u}(\vartheta_0), Y_{i,j,v}(\vartheta_0)), \quad (17)$$

wobei

$$\text{cov}_{\vartheta_0}(Y_{i,j,u}(\vartheta_0), Y_{i,j,v}(\vartheta_0)) = \frac{\Delta b_u(t_{i,n}, \vartheta_0) \Delta b_v(t_{i,n}, \vartheta_0)}{2(\Delta b(t_{i,n}, \vartheta_0))^2} + \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)}. \quad (18)$$

Wir untersuchen jetzt den Stetigkeitsmodul der zweiten Ableitung der Log - Likelihoodfunktion. Dabei sei K_γ die abgeschlossene Kugel im \mathbb{R}^k mit dem Radius γ und dem Mittelpunkt ϑ_0 .

Lemma 7 Die Funktionen m und b mögen den Voraussetzungen (V3) und (V4) genügen. $\vartheta_0 \in \Theta$ sei ein innerer Punkt. Dann gilt für alle $1 \leq u, v \leq k$

$$\lim_{\gamma \downarrow 0} \sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{k_n m_n} |L_{n,u,v}(\vartheta) - L_{n,u,v}(\vartheta_0)| = 0.$$

Beweis: Die abgeschlossene und beschränkte Menge $G \subset \mathbb{R}^2$ sei definiert durch

$$G = \{(t, h) : 0 \leq t \leq T, 0 \leq h \leq T, 0 \leq t + h \leq T\}.$$

γ_0 sei so klein, daß $K_{\gamma_0} \subset \Theta^0$ gilt. Für $(t, h, \vartheta) \in G \times K_{\gamma_0}$ setzen wir

$$\Phi(t, h, \vartheta) = \begin{cases} \frac{b_u(t+h, \vartheta) - b_u(t, \vartheta)}{b(t+h, \vartheta) - b(t, \vartheta)} & h \neq 0, \\ \frac{\dot{b}_u(t, \vartheta)}{\dot{b}(t, \vartheta)} & h = 0. \end{cases}$$

Aufgrund der Voraussetzungen (V3) und (V4) ist $\Phi(t, h, \vartheta)$ eine stetige Funktion auf der kompakten Menge $G \times K_{\gamma_0}$. Damit ist die Folge der Funktionen

$$\vartheta \rightarrow \frac{b_u(t_{i,n}, \vartheta) - b_u(t_{i-1,n}, \vartheta)}{b(t_{i,n}, \vartheta) - b(t_{i-1,n}, \vartheta)}$$

auf K_{γ_0} gleichmäßig beschränkt und gleichgradig stetig. Analog zeigt man, daß die weiteren in der Darstellung (16) von $L_{j,n,u,v}(\vartheta)$ auftretenden Funktionenfolgen, bei denen gewisse Potenzen von $\Delta b(t_{i,n}, \vartheta)$ im Nenner stehen, ebenfalls die Eigenschaften besitzen, gleichmäßig beschränkt und gleichgradig stetig zu sein. Damit erhalten wir die Darstellung

$$\begin{aligned} L_{j,n,u,v} &= \sum_{i=1}^{k_n} \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta))^2}{\Delta b(t_{i,n}, \vartheta)} \varphi_{i,n}(\vartheta) \\ &+ \sum_{i=1}^{k_n} \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}, \vartheta))^{1/2}} \psi_{i,n}(\vartheta) + \sum_{i=1}^{k_n} \rho_{i,n}(\vartheta) = L_n^{(1)}(\vartheta) + L_n^{(2)}(\vartheta) + L_n^{(3)}(\vartheta) \end{aligned} \quad (19)$$

mit auf K_{γ_0} gleichgradig stetigen Familien von Funktionen $\varphi_{i,n}(\vartheta)$, $\psi_{i,n}(\vartheta)$ und $\rho_{i,n}(\vartheta)$. Es gilt

$$\begin{aligned} &|L_n^{(1)}(\vartheta) - L_n^{(1)}(\vartheta_0)| \\ &\leq \sum_{i=1}^{k_n} \left| \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta))^2}{\Delta b(t_{i,n}, \vartheta)} - \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta_0))^2}{\Delta b(t_{i,n}, \vartheta_0)} \right| |\varphi_{i,n}(\vartheta)| \\ &+ \sum_{i=1}^{k_n} \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta_0))^2}{\Delta b(t_{i,n}, \vartheta_0)} |\varphi_{i,n}(\vartheta) - \varphi_{i,n}(\vartheta_0)| \\ &= B_{1,n}(\vartheta) + B_{2,n}(\vartheta). \end{aligned}$$

Weiter gilt

$$\begin{aligned}
 B_{1,n}(\vartheta) &\leq \sum_{i=1}^{k_n} \frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta_0))^2}{\Delta b(t_{i,n}, \vartheta_0)} \cdot \left| \frac{\Delta b(t_{i,n}, \vartheta) - \Delta b(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta)} \varphi_{i,n}(\vartheta) \right| \\
 &+ 2 \sum_{i=1}^{k_n} \left| \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} \right| \cdot \left| (\Delta m(t_{i,n}, \vartheta) - \Delta m(t_{i,n}, \vartheta_0)) \cdot \frac{\Delta b(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta)} \right| |\varphi_{i,n}(\vartheta)| \\
 &+ \sum_{i=1}^{k_n} \frac{|\Delta m(t_{i,n}, \vartheta) - \Delta m(t_{i,n}, \vartheta_0)|^2}{\Delta b(t_{i,n}, \vartheta)} |\varphi_{i,n}(\vartheta)|.
 \end{aligned}$$

Durch ähnliche Überlegungen wie beim Beginn des Beweises läßt sich zeigen, daß die in Betragsstrichen stehenden Familien von Funktionen von ϑ auf K_{γ_0} gleichgradig stetig sind und für $\vartheta = \vartheta_0$ den Wert Null annehmen. Somit existiert eine Funktion $w(\gamma) \rightarrow 0$, $\gamma \downarrow 0$, derart, daß für $\gamma \leq \gamma_0$ gilt

$$\sup_{\vartheta \in K_\gamma} \frac{1}{k_n} B_{1,n}(\vartheta) \leq w(\gamma) \frac{1}{k_n} \sum_{i=1}^{k_n} \left(\frac{(X_{i,j} - \Delta m(t_{i,n}, \vartheta_0))^2}{\Delta b(t_{i,n}, \vartheta_0)} + 2 \left| \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} \right| + 1 \right).$$

Aus $\mathcal{L}_{\vartheta_0} \left(\frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}, \vartheta_0))^{1/2}} \right) = N(0, 1)$ folgt

$$\sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{k_n} B_{1,n}(\vartheta) \leq 4w(\gamma) \xrightarrow{\gamma \rightarrow 0} 0.$$

Analog ergibt sich mit Hilfe der gleichgradigen Stetigkeit der $\varphi_{i,n}(\vartheta)$ die Beziehung

$$\sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{k_n} B_{2,n}(\vartheta) \xrightarrow{\gamma \rightarrow 0} 0.$$

Damit folgt

$$\sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \left| L_n^{(1)}(\vartheta) - L_n^{(1)}(\vartheta_0) \right| \xrightarrow{\gamma \rightarrow 0} 0.$$

Mit ähnlichen Methoden lassen sich analoge Aussagen für $L_n^{(2)}(\vartheta)$ und $L_n^{(3)}(\vartheta)$ nachweisen. Die Behauptung von Lemma 7 folgt jetzt aus der Darstellung (14). ■

Wir wollen nun eine zu Lemma 7 analoge Aussage für die Situation in Theorem 2 nachweisen.

Lemma 8 Die Funktionen b und m mögen den Voraussetzungen in Theorem 2 genügen. $\vartheta_0 \in \Theta$ sei ein innerer Punkt. Dann gilt für alle $\frac{1}{2} \leq u, v \leq k$

$$\lim_{\gamma \downarrow 0} \sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{m_n} |L_{n,u,v}(\vartheta) - L_{n,u,v}(\vartheta_0)| = 0.$$

Beweis: Da b von ϑ unabhängig ist, vereinfacht sich die Darstellung von $L_{j,n,u,v}(\vartheta)$ in (16) zu

$$\begin{aligned}
 L_{j,n,u,v}(\vartheta) &= \sum_{i=1}^{k_n} \left[\frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}))^{1/2}} \cdot \frac{\Delta m_{u,v}(t_{i,n}, \vartheta)}{(\Delta b(t_{i,n}))^{1/2}} - \frac{\Delta m_u(t_{i,n}, \vartheta) \Delta m_v(t_{i,n}, \vartheta)}{\Delta b(t_{i,n})} \right] \\
 &= L_{1,j,n}(\vartheta) + L_{2,j,n}(\vartheta).
 \end{aligned}$$

Es gilt mit $\phi_{i,n}(\vartheta) = \frac{\Delta m_{u,v}(t_{i,n}, \vartheta)}{\Delta b(t_{i,n})}$

$$\begin{aligned} L_{1,j,n}(\vartheta) - L_{1,j,n}(\vartheta_0) &= \sum_{i=1}^{k_n} \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}} (\phi_{i,n}(\vartheta) - \phi_{i,n}(\vartheta_0)) (\Delta b(t_{i,n}))^{1/2} \\ &+ \sum_{i=1}^{k_n} \frac{\Delta m(t_{i,n}, \vartheta) - \Delta m(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} \phi_{i,n}(\vartheta) \Delta b(t_{i,n}). \end{aligned}$$

Wir untersuchen die erste Summe. Weil $\dot{m}_{u,v}$ einer Höderbedingung in einer Umgebung von ϑ_0 genügt, gibt es eine Konstante c und ein $\alpha > 0$ und ein hinreichend kleines γ_0 mit

$$|\phi_{i,n}(\vartheta_1) - \phi_{i,n}(\vartheta_2)| \leq c_1 \|\vartheta_1 - \vartheta_2\|^\alpha$$

für alle $\vartheta_i \in K_{\gamma_0}$. Bezüglich $P_{\vartheta_0}^n$ sind die $\frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}}$ standardnormalverteilt. Deshalb besitzt

$$\xi_{j,n}(\vartheta) = \sum_{i=1}^{k_n} \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}} (\phi_{i,n}(\vartheta) - \phi_{i,n}(\vartheta_0)) (\Delta b(t_{i,n}))^{1/2}$$

eine Normalverteilung mit dem Erwartungswert 0 und der Varianz

$$\sigma_n^2(\vartheta) = \sum_{i=1}^{k_n} (\phi_{i,n}(\vartheta) - \phi_{i,n}(\vartheta_0))^2 \Delta b(t_{i,n}).$$

Es gilt $E(\xi_{j,n}(\vartheta))^{2l} = 1 \cdot 3 \cdots (2l-1) \sigma^{2l}$. Aus dem Mittelwertsatz der Differentialrechnung folgt $\Delta b(t_{i,n}) \leq c_3(t_{i,n} - t_{i-1,n})$ mit einer gewissen Konstanten c_3 . Damit gilt

$$E(\xi_{j,n}(\vartheta))^{2l} \leq 1 \cdot 3 \cdots (2l-1) (c_1^2 \gamma_0^{2\alpha} c_3 T)^l.$$

Analog ergibt sich

$$E(\xi_{j,n}(\vartheta) - \xi_{j,n}(\vartheta_0))^{2l} \leq 1 \cdot 3 \cdots (2l-1) (c_1^2 \|\vartheta_1 - \vartheta_2\|^{2\alpha} c_3 T)^l.$$

Wählt man jetzt l so groß, daß $r = 2\alpha l > k$ gilt, so sind mit $l = 2m$ die Voraussetzungen von Lemma 1 erfüllt, woraus

$$\limsup_{\gamma \downarrow 0} \sup_n E_{\vartheta_0} \sup_{\substack{j,n \\ \vartheta \in K_\gamma}} |\xi_{j,n}(\vartheta) - \xi_{j,n}(\vartheta_0)| = 0$$

folgt. Wir untersuchen jetzt den zweiten Term in der Darstellung von $L_{1,j,n}(\vartheta) - L_{1,j,n}(\vartheta_0)$.

Die Folge $\phi_{i,n}(\vartheta)$ ist gleichmäßig beschränkt auf K_{γ_0} und die Folge $\frac{\Delta m(t_{i,n}, \vartheta) - \Delta m(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})}$

ist gleichgradig stetig. Aus $\Delta b(t_{i,n}) \leq c_3(t_{i,n} - t_{i-1,n})$ ergibt sich dann

$$\limsup_{\gamma \downarrow 0} \sup_{\substack{j,n \\ \vartheta \in K_\gamma}} \left| \sum_{i=1}^{k_n} \frac{\Delta m(t_{i,n}, \vartheta) - \Delta m(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} \phi_{i,n}(\vartheta) \Delta b(t_{i,n}) \right| = 0.$$

Somit gilt

$$\lim_{\gamma \downarrow 0} \sup_{\substack{j,n \\ \vartheta \in K_\gamma}} |L_{1,j,n}(\vartheta) - L_{1,j,n}(\vartheta_0)| = 0.$$

Analog zeigt man

$$\lim_{\gamma \downarrow 0} \sup_{\substack{j,n \\ \vartheta \in K_\gamma}} |L_{2,j,n}(\vartheta) - L_{2,j,n}(\vartheta_0)| = 0.$$

Damit ist Lemma 8 vollständig bewiesen. ■

Beweis von Theorem 1: Weil ϑ_0 ein innerer Punkt von Θ ist, gibt es ein $\gamma_0 > 0$ derart, daß $K_{\gamma_0} \subseteq \Theta^0$ gilt. Im weiteren bezeichnen wir durch $o_{P_{\vartheta_0}^n}(1)$ Folgen von Zufallsvariablen, die $P_{\vartheta_0}^n$ -stochastisch gegen Null streben. Wir setzen $A_n = \{\widehat{\vartheta}_n \in K_{\gamma_0}\}$. Weil $\widehat{\vartheta}_n$ konsistent ist, gilt

$$P_{\vartheta_0}^n(\overline{A_n}) \xrightarrow{n \rightarrow \infty} 0. \quad (20)$$

Auf dem Ereignis A_n besitzt die Log - Likelihoodfunktion ein relatives Maximum in $\widehat{\vartheta}_n$. Deshalb gilt auf A_n für alle u ($u = 1, \dots, k$) $L_{n,u}(\widehat{\vartheta}_n) = 0$. Aufgrund der Beziehung (20) ist $I_{\overline{A_n}} L_{n,u}(\widehat{\vartheta}_n) = o_{P_{\vartheta_0}^n}(1)$.

I_A ist hierbei die Indikatorfunktion des Ereignisses A . Wir bezeichnen durch $DL_n(\vartheta)$ den Gradienten von L_n und durch D^2 die Matrix der zweiten Ableitungen. Durch Taylorentwicklung ergibt sich aus $L_{n,u}(\widehat{\vartheta}_n) = 0$

$$o_{P_{\vartheta_0}^n}(1) = \frac{1}{\sqrt{k_n m_n}} DL_n(\vartheta_0) + \sqrt{k_n m_n} (\widehat{\vartheta}_n - \vartheta_0) \cdot \frac{1}{k_n m_n} (D^2 L_n(\vartheta_0) + R_n), \quad (21)$$

wobei $R_n = D^2 L_n(\widetilde{\vartheta}_n) - D^2 L_n(\vartheta_0)$ gilt und $\widetilde{\vartheta}_n$ auf der Verbindungsgeraden von ϑ_0 und $\widehat{\vartheta}_n$ zwischen diesen beiden Punkten liegt.

Wir untersuchen zunächst R_n . Aus der Konsistenz von $\widehat{\vartheta}_n$ folgt

$$P_{\vartheta_0}^n(\widehat{\vartheta}_n \notin K_\gamma) \xrightarrow{n \rightarrow \infty} 0$$

für alle $\gamma > 0$. Wir schätzen die Elemente $L_{n,u,v}(\widetilde{\vartheta}_n)$ der Matrix R_n ab:

$$\begin{aligned} & P_{\vartheta_0}^n \left(\frac{1}{k_n m_n} |L_{n,u,v}(\widetilde{\vartheta}_n) - L_{n,u,v}(\vartheta_0)| > \varepsilon \right) \\ & \leq P_{\vartheta_0}^n \left(\frac{1}{k_n m_n} |L_{n,u,v}(\widetilde{\vartheta}_n) - L_{n,u,v}(\vartheta_0)| > \varepsilon, \widetilde{\vartheta}_n \in K_\gamma \right) + P_{\vartheta_0}^n(\widetilde{\vartheta}_n \notin K_\gamma) \\ & \leq \frac{1}{\varepsilon} \sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{k_n m_n} |L_{n,u,v}(\vartheta) - L_{n,u,v}(\vartheta_0)| + P_{\vartheta_0}^n(\widetilde{\vartheta}_n \notin K_\gamma). \end{aligned}$$

Wenden wir Lemma 7 auf die rechte Seite an, so folgt

$$\begin{aligned} & \overline{\lim}_{n \rightarrow \infty} P_{\vartheta_0}^n \left(\frac{1}{k_n m_n} |L_{n,u,v}(\tilde{\vartheta}_n) - L_{n,u,v}(\vartheta_0)| > \varepsilon \right) \\ & \leq \frac{1}{\varepsilon} \limsup_{\gamma > 0} \sup_n E_{\vartheta_0} \sup_{\vartheta \in K_\gamma} \frac{1}{k_n m_n} |L_{n,u,v}(\vartheta) - L_{n,u,v}(\vartheta_0)| + \lim_{\gamma > 0} \overline{\lim}_{n \rightarrow \infty} P_{\vartheta_0}^n (\tilde{\vartheta}_n \notin K_\gamma) = 0. \end{aligned}$$

Damit gilt

$$\frac{1}{k_n m_n} R_n = o_{P_{\vartheta_0}^n}(1). \quad (22)$$

Wir untersuchen jetzt die Elemente der Matrix $\frac{1}{k_n m_n} D^2 L_n(\vartheta_0)$. Sei

$$Y_{i,j,n} = \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}, \vartheta_0))^{1/2}}.$$

Aus (14) und (16) folgt bei festem u, v mit

$$S_{i,n} = \frac{\Delta b_{u,v}(t_{i,n}, \vartheta_0)}{2\Delta b(t_{i,n}, \vartheta_0)} - \frac{\Delta b_u(t_{i,n}, \vartheta_0) \Delta b_v(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}, \vartheta_0))^2}$$

$$\begin{aligned} \frac{1}{k_n m_n} L_{n,u,v}(\vartheta_0) &= \frac{1}{m_n} \sum_{j=1}^{m_n} \frac{1}{k_n} \sum_{i=1}^{k_n} (Y_{i,j,n}^2 - 1) S_{i,n} + \frac{1}{m_n} \sum_{j=1}^{m_n} \frac{1}{k_n} \sum_{i=1}^{k_n} Y_{i,j,n} \psi_{i,n}(\vartheta_0) \\ &\quad - \frac{1}{m_n} \sum_{j=1}^{m_n} \frac{1}{k_n} \sum_{i=1}^{k_n} \left[\frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} + \frac{\Delta b_u(t_{i,n}, \vartheta_0) \Delta b_v(t_{i,n}, \vartheta_0)}{2(\Delta b(t_{i,n}, \vartheta_0))^2} \right], \end{aligned} \quad (23)$$

wobei die $S_{i,n}$ und $\psi_{i,n}(\vartheta_0)$ gleichmäßig beschränkt sind. Wegen $\mathcal{L}_{\vartheta_0}(Y_{i,j,n}) = N(0,1)$ und Lemma 6 folgt

$$E_{\vartheta_0} \left(\frac{1}{k_n} \sum_{i=1}^{k_n} (Y_{i,j,n}^2 - 1) S_{i,n} \right)^2 \xrightarrow{n \rightarrow \infty} 0 \quad \text{und} \quad E_{\vartheta_0} \left(\frac{1}{k_n} \sum_{i=1}^{k_n} Y_{i,j,n} \psi_{i,n}(\vartheta_0) \right)^2 \xrightarrow{n \rightarrow \infty} 0.$$

Damit streben die ersten beiden Terme auf der rechten Seite von (23) $P_{\vartheta_0}^n$ -stochastisch gegen Null für $n \rightarrow \infty$. Wegen Lemma 3 strebt der dritte Term für $n \rightarrow \infty$ gegen $-\sigma_{u,v}(\vartheta_0)$. Damit strebt also $\frac{1}{k_n m_n} (D^2 L_n(\vartheta_0) + R_n) P_{\vartheta_0}^n$ -stochastisch gegen $-\Sigma(\vartheta_0)$.

Wir untersuchen nun die Folge der zufälligen Vektoren $\frac{1}{\sqrt{k_n m_n}} D L_n(\vartheta_0)$. $\langle \cdot, \cdot \rangle$ bezeichne das Skalarprodukt in \mathbb{R}^k und $a \in \mathbb{R}^k$ sei fest gewählt. Dann gilt aufgrund von (16)

$$\left\langle \frac{1}{\sqrt{k_n m_n}} D L_n(\vartheta_0), a \right\rangle = \frac{1}{\sqrt{k_n m_n}} \sum_{j=1}^{m_n} \sum_{i=1}^{k_n} \left((Y_{i,j,n}^2 - 1) c_{i,n} + Y_{i,j,n} d_{i,n} \right),$$

wobei

$$c_{i,n} = \frac{1}{2} \sum_{u=1}^k \frac{\Delta b_u(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n}, \vartheta_0)} a_u \quad \text{und} \quad d_{i,n} = \sum_{u=1}^k \frac{\Delta m_u(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}, \vartheta_0))^{1/2}} a_u.$$

Aus Lemma 3 folgt

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{k_n} \sum_{i=1}^{k_n} (2c_{i,n}^2 + d_{i,n}^2) = \sum_{u,v=1}^k \sigma_{u,v}(\vartheta_0) a_u a_v.$$

Wenden wir jetzt Lemma 5 an, so folgt

$$\mathcal{L}_{\vartheta_0} \left(\left\langle \frac{1}{k_n m_n} D L_n(\vartheta_0), a \right\rangle \right) \Rightarrow N(0, \sigma^2).$$

Durch Anwendung der Cramer - Wold - Technik erhält man

$$\mathcal{L}_{\vartheta_0} \left(\frac{1}{\sqrt{k_n m_n}} D L_n(\vartheta_0) \right) \Rightarrow N(0, \Sigma(\vartheta_0)).$$

Weil $\frac{1}{k_n m_n} (D^2 L_n(\vartheta_0) + R_n)$ stochastisch gegen die reguläre Matrix $\Sigma(\vartheta_0)$ strebt, ergibt sich die Behauptung aus der Darstellung (21) durch Anwendung des Satzes von Slutsky. ■

Beweis von Theorem 2:

Durch Anwendung der gleichen Argumente wie in Beweis von Theorem 1 erhält man

$$o_{P_{\vartheta_0}^n}(1) = \frac{1}{\sqrt{m_n}} D L_n(\vartheta_0) + \frac{1}{m_n} \left(D^2 L_n(\vartheta_0) \sqrt{m_n} (\widehat{\vartheta}_n - \vartheta_0) + R_n \right).$$

Wegen Lemma 8 strebt die Folge der Matrizen $\frac{1}{m_n} R_n$ $P_{\vartheta_0}^n$ -stochastisch gegen Null. Wir setzen

$$Y_{i,j,n} = \frac{X_{i,j} - \Delta m(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}}.$$

Die u -te Komponente von $\frac{1}{\sqrt{m_n}} D L_n(\vartheta_0)$ hat dann die Gestalt

$$\frac{1}{\sqrt{m_n}} L_{n,u}(\vartheta_0) = \frac{1}{\sqrt{m_n}} \sum_{j=1}^{m_n} \sum_{i=1}^{k_n} Y_{i,j,n} \frac{\Delta m_u(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}}. \quad (24)$$

Die Elemente der Matrix $\frac{1}{m_n} D^2 L_n(\vartheta_0)$ haben die Form

$$\frac{1}{m_n} L_{n,u,v}(\vartheta_0) = \frac{1}{m_n} \sum_{j=1}^{m_n} \left[\sum_{i=1}^{k_n} \left(Y_{i,j,n} \frac{\Delta m_{u,v}(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}} - \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} \right) \right]. \quad (25)$$

Wir setzen $Z_{i,n} = \frac{1}{m_n} \sum_{j=1}^{m_n} Y_{i,j,n}$. Dann gilt $\mathcal{L}_{\vartheta_0}(Z_{i,n}) = N\left(0, \frac{1}{m_n}\right)$. Wegen Voraussetzung

(V4) ist $\frac{\Delta b_{u,v}(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})}$, $i = 1, \dots, k_n$, $n = 1, \dots$, betragsmäßig beschränkt. Wegen der Voraussetzung über b in Theorem 2 gilt $\Delta b(t_{i,n}) \leq \sup_{0 \leq t \leq T} \dot{b}(t)(t_{i,n} - t_{i-1,n})$ mit $\sup_{0 \leq t \leq T} \dot{b}(t) < \infty$.

Damit hat der erste Term auf der rechten Seite von (25) die Gestalt $\sum_{i=1}^{k_n} Z_{i,n} h_{i,n}$, wobei die

$h_{i,n}$ mit einer gewissen Konstanten c die Abschätzung $|h_{i,n}| \leq c(t_{i,n} - t_{i-1,n})^{1/2}$ gestatten. Damit ist $U_n = \sum_{i=1}^{k_n} Z_{i,n} h_{i,n}$ eine normalverteilte Zufallsgröße mit dem Erwartungswert Null und einer Varianz, die höchstens $\frac{c^2 T}{m_n}$ ist. Wegen $m_n \rightarrow \infty$ streben deshalb die $U_n P_{\vartheta_0}^n$ -stochastisch gegen Null. Da der zweite Term in (25) nach Lemma 5 gegen $-\widetilde{\sigma}_{u,v}$ strebt, konvergiert insgesamt $\frac{1}{m_n} D^2 L_n(\vartheta_0) P_{\vartheta_0}^n$ -stochastisch gegen $-\widetilde{\Sigma}(\vartheta_0)$.

Wie im Beweis von Theorem 1 bleibt nur zu zeigen, daß die Verteilung von $\frac{1}{\sqrt{m_n}} D L_n(\vartheta_0)$ gegen $N(0, \widetilde{\Sigma}(\vartheta_0))$ strebt. Diese Aussage läßt sich folgendermaßen begründen. Die $U_{i,n} = \frac{1}{\sqrt{m_n}} \sum_{j=1}^{m_n} Y_{i,j,n}$ sind unabhängige standardnormalverteilte Zufallsvariable. Somit besitzt der Vektor

$$\left(\sum_{i=1}^{k_n} U_{i,n} \frac{\Delta m_1(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}}, \dots, \sum_{i=1}^{k_n} U_{i,n} \frac{\Delta m_k(t_{i,n}, \vartheta_0)}{(\Delta b(t_{i,n}))^{1/2}} \right)$$

eine k -dimensionale Normalverteilung mit Erwartungswertvektor 0 und der Kovarianzmatrix

$$\Sigma_n = \left(\sum_{i=1}^{k_n} \frac{\Delta m_u(t_{i,n}, \vartheta_0) \Delta m_v(t_{i,n}, \vartheta_0)}{\Delta b(t_{i,n})} \right)_{1 \leq u, v \leq k}$$

Wegen Lemma 4 gilt $\Sigma_n \xrightarrow{n \rightarrow \infty} \widetilde{\Sigma}(\vartheta_0)$ und hieraus folgt

$$\mathcal{L}_{\vartheta_0} \left(\frac{1}{\sqrt{m_n}} D L_n(\vartheta_0) \right) \Rightarrow N \left(0, \widetilde{\Sigma}(\vartheta_0) \right). \quad \blacksquare$$

Literaturverzeichnis

- [1] Ibragimov, I.A. and Hasminskij, R.Z. : *Asymptotische Schätztheorie*. (Russisch) Moskau 1979
- [2] Jacod, J. : *Random sampling in estimation problems for continuous Gaussian processes with independent increments*. Preprint No. 57, Laboratoire de Probabilités de L'Université Paris V, 1991
- [3] Jacod, J. and Shiryaev, A. N. : *Limit Theorems for Stochastic Processes*. Berlin 1987
- [4] Liese, F. and Wienke, A. : *Exponential rate of convergence of maximum likelihood estimators for inhomogeneous Wiener processes*. Eingereicht bei Kybernetika (Prague)

eingegangen: 2. Dezember 1992

Autoren:

Prof. Dr. Friedrich Liese
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Bundesrepublik Deutschland

Dipl.-Math. Andreas Wienke
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Bundesrepublik Deutschland

WALTRAUT KAHLE; FRIEDRICH LIESE

Consistency and Asymptotic Normality of Minimum Contrast Estimation in Renewal Processes

ABSTRACT. The consistency and asymptotic normality of minimum contrast estimation (which includes the maximum likelihood estimation as a special case) is established if the sample is from a renewal process and the observation time tends to infinity. It is shown, that the conditions for consistency and asymptotic normality for maximum likelihood estimation are fulfilled if the distribution of the time between two renewals is Weibull or Maxwell. This fact is used to construct simultaneous confidence regions for the parameters of the Weibull distribution and a confidence interval for the parameter of the Maxwell distribution.

1 Introduction

In reliability theory and statistical quality control very often renewal processes are observed. If the distribution of the lifetime is exponential it is possible to calculate the distribution of the point estimation of the unknown parameter and to derive confidence estimations from this distribution (MANN, SCHAFER, SINGPURWALLA [4]). In other cases asymptotic methods have to be used to calculate confidence estimations of the unknown parameters. There are two kinds of asymptotics:

1. The number m of realizations of renewal processes tends to infinity and each realization is observed a fixed time T_i , $i = 1, \dots, m$.
2. The observation time of at least one realization of the renewal process tends to infinity and the number of realizations is fixed.

In the first case the consistency and asymptotic normality of maximum likelihood estimations can be proved, using theorems for independent, not necessary identically distributed observations (see a.e. BRAGIMOV, HASMINSKI [2]). In the second case we have neither independent nor identically distributed observations and it is impossible to prove the consistency and asymptotic normality of maximum likelihood estimations by classical methods. In this paper conditions which guarantee consistency and asymptotic normality of minimum

contrast estimations are established if the observation time of renewal processes tends to infinity.

It turns out that the asymptotic behaviour of minimum contrast estimators for renewal processes is closely related to the behaviour of the same estimators in the i.i.d. case. The difference is that in our case the number of observations is random. We do not follow PFANZAGL [5] in section 2 where the consistency is established. We require continuity instead of semicontinuity of the contrast function and prove weak consistency. This leads to an essential simplification of the proof. Section 3 contains the proofs of the theorems of section 2. In section 4 the conditions of consistency and asymptotic normality are established for the special case of maximum likelihood estimation if the underlying distribution is Weibull. It is demonstrated, how the asymptotic normality can be used to construct simultaneous confidence regions for the parameters of the Weibull distribution. For the parameter of the Maxwell distribution maximum likelihood estimation and another minimum contrast estimation are compared.

2 Models and results

Suppose X_1, X_2, \dots are i.i.d. nonnegative random variables with distribution function F_θ , where the parameter θ belongs to the parameter set Θ being a subset of R^p . Introduce the renewal process N_t by

$$N_t = \max \left\{ k : \sum_{i=1}^k X_i \leq t \right\}.$$

If $F_\theta(0) < 1$ then $N_t < \infty$. The rest time R_t is defined by

$$R_t = t - \sum_{i=1}^{N_t} X_i.$$

The problem now is to estimate the parameter θ using the sample X_1, \dots, X_{N_t}, R_t .

Often, not only one renewal process but m independent processes are observed with different observation times. More precisely let $X_{1,j}, X_{2,j}, \dots, j = 1, \dots, m$ i.i.d. random variables with distribution function F_θ . Let T_1, \dots, T_m be the observation times and denote by

$$R_{T_j} = T_j - \sum_{i=1}^{N_{T_j}} X_{i,j}$$

the corresponding rest times. Our available information consists in the sample

$$X_{1,j}, \dots, X_{N_{T_j},j}, R_{T_j}, \quad j = 1, \dots, m.$$

To estimate the parameter θ we use the concept of minimum contrast estimations intensively studied in a series of papers by PFANZAGL, a.e. [5] not for renewal processes but for i.i.d. samples.

Let $(P_\theta)_{\theta \in \Theta}$ be a family of distributions on (R, \mathcal{B}) where R is the real line and \mathcal{B} the σ -algebra of Borel sets. Let $\varrho : R \times \Theta \rightarrow [-\infty, \infty]$ be a function so that $\varrho(\cdot, \theta)$ is measurable for every $\theta \in \Theta$. ϱ is called a contrast function if for every $\theta_1, \theta_2 \in \Theta$ the integral $\int \varrho(x, \theta_1) P_{\theta_2}(dx)$ exists and

$$\int \varrho(x, \theta_1) P_{\theta_2}(dx) > \int \varrho(x, \theta_2) P_{\theta_2}(dx) \quad \text{if } \theta_1 \neq \theta_2. \quad (1)$$

Example 1: Suppose P_θ is dominated by a σ -finite measure λ . Denote by $f_\theta = \frac{dP_\theta}{d\lambda}$ the Radon-Nikodym derivative and put

$$\varrho(x, \theta) = -\log f_\theta(x).$$

Then ϱ is a contrast function provided that $\int (\log f_{\theta_2}) dP_{\theta_2} > -\infty$ and $\int (\log f_{\theta_1}) dP_{\theta_2} < \infty$ for every $\theta_1, \theta_2 \in \Theta$ (see PFANZAGL [5]).

Example 2: Suppose $\Theta = (0, \infty)$, μ is the Lebesgue measure and $f(x)$ is a density with respect to μ . Assume $f(x) = 0$ for $x < 0$ and

$$\int_0^\infty x f(x) dx = 1.$$

Introduce a family of densities by $f_\theta(x) = \theta f(\theta x)$. Then obviously $\frac{1}{\theta}$ is the expectation of f_θ . Set

$$\varrho(x, \theta) = -\log \theta + x\theta.$$

Clearly $-\varrho$ is the loglikelihood of the exponential distribution. But we will show that even ϱ is a contrast function for the distribution P_θ with density f_θ . This contrast function will be applied to construct consistent and asymptotically normally distributed estimators.

$$\begin{aligned} & \int (\varrho(x, \theta_1) - \varrho(x, \theta_2)) f_{\theta_2}(x) dx \\ &= \int \left(-\log \frac{\theta_1}{\theta_2} + (\theta_1 - \theta_2)x \right) \theta_2 f(\theta_2 x) dx \\ &= -\log \frac{\theta_1}{\theta_2} + \frac{\theta_1}{\theta_2} - 1 \geq 0. \end{aligned}$$

As $-\log x - 1 + x \geq 0$ with equality iff $x = 1$ we get that ϱ is a contrast function.

To apply the concept of contrast function to renewal processes we introduce

$$Q(\tau) = \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} \varrho(X_{ij}, \tau) + \sum_{j=1}^m \tilde{\varrho}(R_{T_j}, \tau) \quad (2)$$

where ϱ is a contrast function and $\tilde{\varrho}(\cdot, \tau)$ is any measurable function which includes the information contained in the rest time.

X_{ij} , $j = 1, \dots, m$, $i = 1, 2, \dots$ are i.i.d. with common distribution P_θ . Now, roughly speaking, the minimum contrast estimation is defined by minimization of $Q(\tau)$ over $\tau \in \Theta$. To be more precise we denote by $\Omega = R \times R \times \dots$ the set of all sequences of real numbers and by \mathcal{F} Kolmogorov's σ -algebra. Write $(\Omega, \mathcal{F})^m$ for the m -times product space. Suppose $\Theta \subseteq R^p$ is a Borel set and denote by \mathcal{B}_Θ the σ -algebra of Borel subsets of Θ . By \mathcal{E} we shall mean the class of all measurable mappings

$$S : (\Omega, \mathcal{F})^m \longrightarrow (\Theta, \mathcal{B}_\Theta)$$

which depends only on the observations up to the times T_1, \dots, T_j , i.e. which may be represented as

$$S = \sum_{k_1, \dots, k_m=0}^{\infty} \varphi_{k_1, \dots, k_m}(X_{1,1}, \dots, X_{k_m, m}, R_{T_1}, \dots, R_{T_m}) I_{A_{k_1, \dots, k_m}}$$

where

$$\varphi_{k_1, \dots, k_m} : R^{N+m} \longrightarrow \Theta, \quad N = \sum_{i=1}^m k_i$$

are measurable functions and $I_{A_{k_1, \dots, k_m}}$ is the indicator function of the event

$$A_{k_1, \dots, k_m} = \{N_{T_1} = k_1, \dots, N_{T_m} = k_m\}.$$

Now a minimum contrast estimator $\hat{\theta}$ is an estimator $\hat{\theta}$ which belongs to \mathcal{E} and minimizes $Q(\theta)$ over Θ .

Suppose $P(X_{i,j} > 0) > 0$. Then $N_{T_i} < \infty$ a.s. Assume further that Θ is compact and both $\tau \rightarrow \varrho(x, \tau)$ and $\tau \rightarrow \tilde{\varrho}(x, \tau)$ are lower semicontinuous for every $x \in R$. In this situation one can easily shown by similar methods as in PFANZAGL [5] that a measurable minimum contrast estimation exists.

Assume the distributions P_θ of $X_{i,j}$, $\theta \in \Theta$, are measures theoretical equivalent and put for any fixed $\theta_0 \in \Theta$ $f_\theta = \frac{dP_\theta}{dP_{\theta_0}}$. If F_θ (distribution function of $X_{i,j}$) is continuous, then the loglikelihood of the j -th renewal process in $[0, T_j]$ is given by [3]:

$$\log \frac{dP_\theta^{T_j}}{dP_{\theta_0}^{T_j}} = \sum_{i=1}^{N_{T_j}} \log f_\theta(x_{i,j}) + \log \frac{1 - F_\theta(R_{T_j})}{1 - F_{\theta_0}(R_{T_j})} \quad (3)$$

where $P_\theta^{T_j}$ and $P_{\theta_0}^{T_j}$ are the distributions of the renewal processes on $[0, T_j]$ with parameter θ and θ_0 , respectively. Comparing (2) and (3) it is easy to see that the maximum likelihood estimation is a special minimum contrast estimation.

Assume $T_{j,1} \leq T_{j,2} \leq \dots$ for $j = 1, \dots, m$ and put $T_n = \sum_{j=1}^m T_{j,n}$.

Theorem 1 Suppose Θ is compact and both $\varrho(x, \theta)$ and $\tilde{\varrho}(x, \theta)$ are continuous functions on $[0, \infty) \times \Theta$ with values in $(-\infty, +\infty)$. Assume $(P_\theta)_{\theta \in \Theta}$ is the distribution of the nonnegative i.i.d. $X_{i,j}$ and ϱ is a contrast function with respect to $(P_\theta)_{\theta \in \Theta}$. Suppose

$$\int \sup_{\theta \in \Theta} |\varrho(x, \theta)| P_\theta(dx) < \infty.$$

If $0 < \mu = EX_{i,j} < \infty$ and $T_n = \sum_{j=1}^m T_{j,n} \rightarrow \infty$ for $n \rightarrow \infty$ then every minimum contrast estimation $\hat{\theta}_n$ which minimizes (2) is weakly consistent, i.e. for every $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| > \varepsilon) = 0.$$

We now establish conditions for asymptotic normality of minimum contrast estimates. Assume that $\lim_{n \rightarrow \infty} \frac{T_{j,n}}{T_n} = \alpha_j$ exists.

Theorem 2 Assume the conditions of Theorem 1 are fulfilled. Suppose in addition that the true parameter θ belongs to the interior Θ^0 of Θ and $\tilde{\varrho}(x, \theta)$, $\varrho(x, \theta)$ are twice continuously differentiable with respect to each θ_j , $j = 1, \dots, p$, $\theta = (\theta_1, \dots, \theta_p)$ in Θ^0 and there exists a neighborhood $U_\theta \subseteq \Theta^0$ of θ so that

$$\int \sup_{\tau \in U_\theta} M(\tau) P_\theta(dx) < \infty \quad (4)$$

$$\text{with} \quad M(\tau) = \sum_{j,k=1}^p \left| \frac{\partial^2}{\partial \theta_j \partial \theta_k} \varrho(x, \tau) \right|$$

and

$$\int \left(\frac{\partial}{\partial \theta_j} \varrho(x, \tau) \right)^2 P_\theta(dx) < \infty. \quad (5)$$

If $\det \Sigma(\theta) \neq 0$ with

$$\Sigma(\theta) = ((\Sigma_{jk})), \quad j, k = 1, \dots, p, \quad \Sigma_{jk} = \int \frac{\partial^2}{\partial \theta_j \partial \theta_k} \varrho(x, \theta) P_\theta(dx)$$

and $E \frac{\partial}{\partial \theta_k} \varrho(X_{11}, \theta) = 0$, $k = 1, \dots, p$, then the distribution of

$$\sqrt{T_n} (\hat{\theta}_n - \theta)$$

converges weakly to a p -dimensional normal distribution with expectation zero and covariance matrix $C(\theta) = \frac{1}{\mu} \Sigma^{-1}(\theta) \tilde{\Sigma}(\theta) \Sigma^{-1}(\theta)$ where

$$\tilde{\Sigma}(\theta) = \left(\int \frac{\partial}{\partial \theta_i} \varrho(x, \theta) \frac{\partial}{\partial \theta_j} \varrho(x, \theta) P_\theta(dx) \right)_{1 \leq i, j \leq p}.$$

We remark that in the case of $\varrho = -\log f_\theta$ it holds $\Sigma = \tilde{\Sigma}$ and $C = \frac{1}{\mu} \Sigma^{-1} = \frac{1}{\mu} I^{-1}$ with the Fisher information matrix I .

3 Proofs

Lemma 3 Suppose $\Theta \subseteq \mathbb{R}^p$ is compact and $\tau \rightarrow h(X_1, \tau)$ is a.s. continuous. Assume X_1, X_2, \dots are i.i.d. and

$$E \sup_{\tau \in \Theta} |h(X_1, \tau)| < \infty.$$

Then

$$\limsup_{n \rightarrow \infty} \sup_{\tau \in \Theta} \left| \frac{1}{n} \sum_{i=1}^n h(X_i, \tau) - Eh(X_1, \tau) \right| = 0 \quad \text{a.s.}$$

Lemma 3 is a consequence of the strong law of large numbers for random variables taking values in a separable Banach space (see P. REVEZS [6]).

Set $N_T = \max\{k : \sum_{i=1}^k X_i < T\}$.

Lemma 4 Assume the assumptions of Lemma 3 are fulfilled. If $0 < \mu = EX_1 < \infty$ then for every sequence $T_n \rightarrow \infty$

$$\limsup_{T_n \rightarrow \infty} \sup_{\tau \in \Theta} \left| \frac{1}{T_n} \sum_{i=1}^{N_{T_n}} h(X_i, \tau) - \frac{1}{\mu} Eh(X_1, \tau) \right| = 0 \quad \text{a.s.}$$

Proof: Lemma 3 implies

$$\limsup_{n \rightarrow \infty} \sup_{\tau \in \Theta} \left| \frac{1}{N_{T_n}} \sum_{i=1}^{N_{T_n}} h(X_i, \tau) - Eh(X_1, \tau) \right| = 0.$$

The rest follows from

$$\frac{1}{T_n} N_{T_n} \xrightarrow[n \rightarrow \infty]{} \frac{1}{\mu} \quad \text{a.s.}$$

Lemma 5 Let $\{X_{ik}, k = 1, \dots, p, i \geq 1\}$ be a sequence of i.i.d. random vectors with mean 0 and regular covariance matrix Σ . Further, assume that

$$\frac{N_{T_n}}{T_n} \xrightarrow{P} \frac{1}{\mu} \quad (0 < \mu < \infty) \quad \text{as } T_n \rightarrow \infty.$$

Then the distribution of the vectors

$$\left(\frac{1}{\sqrt{N_{T_n}}} \sum_{i=1}^{N_{T_n}} X_{i,k} \right)_{k=1,\dots,p} \quad \text{and} \quad \left(\frac{1}{\sqrt{\mu T_n}} \sum_{i=1}^{N_{T_n}} X_{i,k} \right)_{k=1,\dots,p}$$

converges weakly to a normal distribution with mean zero and covariance matrix Σ as $T_n \rightarrow \infty$.

This lemma may be obtained from ANSCOMBE's theorem using the CRAMER – WALD device (see a.e. GUT [1] and SERVLING [7]).

Proof of Theorem 1: Put

$$Q_{j,n}(\tau) = \frac{1}{T_{j,n}} \left(\sum_{i=1}^{N_{T_{j,n}}} \varrho(X_{i,j}, \tau) + \tilde{\varrho}(R_{T_{j,n}}, \tau) \right)$$

and

$$M_{T_{j,n}} = \inf \left\{ k : \sum_{j=1}^k X_{i,j} > T_{j,n} \right\}.$$

$M_{T_{j,n}}$ is a stopping time with respect to the filtration $\mathcal{F}_{j,n} = \sigma(X_{j,1}, X_{j,2}, \dots)$. Hence by WALD's identity

$$\begin{aligned} E \left[\sum_{i=1}^{N_{T_{j,n}}} \sup_{\tau} |\varrho(X_{i,j}, \tau)| \right] &\leq E \left[\sum_{i=1}^{M_{T_{j,n}}} \sup_{\tau} |\varrho(X_{i,j}, \tau)| \right] \\ &\leq (EM_{T_{j,n}}) E \sup_{\tau} |\varrho(X_{1,1}, \tau)| \\ &\leq (EN_{T_{j,n}} + 1) E \sup_{\tau} |\varrho(X_{1,1}, \tau)|. \end{aligned} \quad (6)$$

Put

$$\frac{1}{T_n} Q_n(\tau) = \sum_{j=1}^m \frac{T_{j,n}}{T_n} Q_{j,n}(\tau).$$

and

$$\varphi(x) = \sup_{\tau} |\tilde{\varrho}(x, \tau)|.$$

By the continuity of $\tilde{\varrho}$ we get $\varphi(x) < \infty$ for every x . As

$$P(R_{T_{j,n}} > a) \leq P(X_{1,1} > a)$$

we see that the sequence $\varphi(R_{T_{j,n}})$ is stochastically bounded. Hence in probability

$$\frac{1}{T_n} \sum_{j=1}^m \varphi(R_{T_{j,n}}) \xrightarrow[n \rightarrow \infty]{} 0. \quad (7)$$

If $\lim_{n \rightarrow \infty} \frac{T_{j,n}}{T_n} = 0$ then by (6)

$$\lim_{n \rightarrow \infty} \sup_{\tau} |Q_{j,n}(\tau)| = 0$$

in probability. If $\lim_{n \rightarrow \infty} \frac{T_{j,n}}{T_n} > 0$ then by Lemma 4 and (7)

$$\lim_{n \rightarrow \infty} P \left(\sup_{\tau} \left| Q_{j,n}(\tau) - \frac{1}{\mu} E \varrho(X_{1,1}, \tau) \right| > \varepsilon \right) = 0.$$

This means

$$\lim_{n \rightarrow \infty} P \left(\sup_{\tau} \left| \frac{1}{T_n} Q_n(\tau) - \frac{1}{\mu} E \varrho(X_{1,1}, \tau) \right| > \varepsilon \right) = 0. \quad (8)$$

As $\varrho(X_{1,1}, \tau)$ depends continuously on τ and $E \sup_{\tau} |\varrho(X_{1,1}, \tau)| < \infty$ the dominated convergence theorem implies that $E \varrho(X_{1,1}, \tau)$ is a continuous function of τ . As ϱ is a contrast function and Θ is compact we get for every $\gamma > 0$

$$\alpha = \inf_{\tau: \|\tau - \theta\| > \gamma} \frac{1}{\mu} E \varrho(X_{1,1}, \tau) - \frac{1}{\mu} E \varrho(X_{1,1}, \theta) > 0.$$

Put for $\beta > 0$

$$A_{\beta} = \left\{ \sup_{\tau} \left| \frac{1}{T_n} Q_n(\tau) - \frac{1}{\mu} E \varrho(X_{1,1}, \tau) \right| \leq \beta \right\}.$$

For τ with $\|\tau - \theta\| > \gamma$ it holds on the event A_{β}

$$\begin{aligned} \frac{1}{T_n} Q_n(\tau) &\geq \frac{1}{\mu} E \varrho(X_{1,1}, \tau) - \beta \\ &\geq \frac{1}{\mu} E \varrho(X_{1,1}, \theta) + \alpha - \beta \\ &\geq \frac{1}{T_n} Q_n(\theta) + \alpha - 2\beta. \end{aligned}$$

Hence

$$\inf_{\|\tau - \theta\| > \gamma} \frac{1}{T_n} Q_n(\tau) \geq \frac{1}{T_n} Q_n(\theta) + \alpha - 2\beta > \frac{1}{T_n} Q_n(\theta) + \frac{2\alpha}{3} - 2\beta$$

and

$$\{\|\hat{\theta}_n - \theta\| > \gamma\} \subseteq \left\{ \inf_{\tau: \|\tau - \theta\| > \gamma} \frac{1}{T_n} Q_n(\tau) > \frac{1}{T_n} Q_n(\theta) \right\} \subseteq A_{\alpha/3}$$

the assertion now follows from (8).

Proof of Theorem 2: Introduce the event A_n by

$$A_n = \{\hat{\theta}_n \in \tilde{U}_{\beta}\}$$

where $\bar{U}_\theta \subseteq U_\theta$ is a closed sphere with centre θ and positive radius. Then by the consistency of $\hat{\theta}_n$

$$P(A_n) \xrightarrow{n \rightarrow \infty} 1.$$

On the event A_n the function $Q_n(\theta)$ attains the minimum at $\hat{\theta}_n$ which is a local minimum. Hence on A_n

$$0 = \sum_{j=1}^m \left(\sum_{i=1}^{N_{T_j,n}} \varrho_k(X_{i,j}, \hat{\theta}_n) + \tilde{\varrho}_k(R_{T_j,n}, \hat{\theta}_n) \right), \quad k = 1, \dots, p \quad (9)$$

where $\varrho_k(x, \theta) = \frac{\partial}{\partial \theta_k} \varrho(x, \theta)$. Similarly, $\varrho_{i,j}$ denotes the derivative with respect to θ_i and θ_j . Taylor expansion and (9) yield

$$\begin{aligned} \alpha_p(1) &= \sum_{j=1}^m \left(\sum_{i=1}^{N_{T_j,n}} \varrho_k(X_{i,j}, \theta) + \tilde{\varrho}_k(R_{T_j,n}, \theta) \right) \\ &+ \sum_{j=1}^m \sum_{l=1}^p (\hat{\theta}_{n,l} - \theta_l) \left[\sum_{i=1}^{N_{T_j,n}} \varrho_{k,l}(X_{i,j}, \bar{\theta}_n) + \tilde{\varrho}_{k,l}(R_{T_j,n}, \bar{\theta}_n) \right] \end{aligned}$$

where $\|\bar{\theta}_n - \theta\| < \|\hat{\theta}_n - \theta\|$. Using the same arguments as in the proof of Theorem 1 we get

$$\frac{1}{\sqrt{T_n}} \sum_{j=1}^m \sup_{\tau \in U_\theta} |\tilde{\varrho}_k(R_{T_j,n}, \tau)|$$

converges stochastically to zero as $n \rightarrow \infty$. Hence

$$\begin{aligned} \alpha_p(1) &= \left(\sum_{j=1}^m \left(\frac{T_{j,n}}{T_n} \right)^{1/2} \cdot \frac{1}{\sqrt{T_{j,n}}} \sum_{i=1}^{N_{T_j,n}} \varrho_k(X_{i,j}, \theta) \right) \\ &+ \sum_{l=1}^p \sqrt{T_n} (\hat{\theta}_{n,l} - \theta_l) \left[\frac{1}{T_n} \sum_{j=1}^m \sum_{i=1}^{N_{T_j,n}} \varrho_{k,l}(X_{i,j}, \bar{\theta}_n) \right]. \end{aligned} \quad (10)$$

In view of Lemma 5 and the assumption $E\varrho_k(X_{1,1}, \theta) = 0$ the distribution of the vectors

$$\left(\frac{1}{\sqrt{\mu T_{j,n}}} \sum_{i=1}^{N_{T_j,n}} \varrho_k(X_{i,j}, \theta) \right)_{k=1, \dots, p}$$

converges weakly to a normal distribution with mean zero and covariance matrix

$$(E\varrho_k(X_{1,1}, \theta)\varrho_l(X_{1,1}, \theta))_{1 \leq k, l \leq p} = \hat{\Sigma}(\theta).$$

By the independence of the m renewal processes we see that the distribution of

$$\sum_{j=1}^m \left(\frac{T_{j,n}}{T_n} \right)^{1/2} \cdot \frac{1}{\sqrt{T_{j,n}}} \sum_{i=1}^{N_{T_j,n}} \varrho_k(X_{i,j}, \theta)$$

weakly converges to a normal distribution with mean zero and covariance matrix $\mu \tilde{\Sigma}(\theta)$. The consistency of $\hat{\theta}_n$ implies the stochastic convergence of $\bar{\theta}_n$ to θ . If we apply the law of large numbers in Lemma 4

$$\frac{1}{T_{j,n}} \sum_{i=1}^{N_{T_j,n}} \varrho_{k,l}(X_{i,j}, \tau), \quad \tau \in \tilde{U}_\theta,$$

we see that

$$\frac{1}{T_n} \sum_{j=1}^m \sum_{i=1}^{N_{T_j,n}} \varrho_{k,l}(X_{i,j}, \bar{\theta}_n)$$

tends stochastically to $\frac{1}{\mu} E \varrho_{k,l}(X_{1,1}, \theta)$ and $\left(\frac{1}{\mu} E \varrho_{k,l}(X_{11}, \theta) \right)_{1 \leq k, l \leq p}$ is the matrix $\frac{1}{\mu} \Sigma(\theta)$.

The rest follows from the representation (10).

4 Consistency and Asymptotic Normality of Minimum Contrast Estimations for the Weibull and the Maxwell Distribution

In many practical problems the Weibull distribution is used to describe the random failure time of an item. In this section the model of section 2 is used to show, that the maximum likelihood estimation of parameters of the Weibull distribution is asymptotically normal distributed, if the sample is the result of observation a renewal process.

In the case of the Weibull distribution the density of renewal time has the form

$$f_\theta(x) = k \cdot \vartheta^{-k} \cdot x^{k-1} \cdot \exp\left(-\left(\frac{x}{\vartheta}\right)^k\right), \quad t > 0. \quad (11)$$

Let be $\theta = (\theta_1, \theta_2) = (k, \vartheta)$, where $k > 0$ is the form parameter and $\vartheta > 0$ is the scale parameter. Θ is assumed to be compact.

From (3) the loglikelihood for observations of m renewal processes in $[0, T_j]$, $j = 1, \dots, m$ is given by

$$\begin{aligned} \log L &= \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} \left(\log k - \log \vartheta + (k-1)(\log x_{i,j} - \log \vartheta) - \left(\frac{x_{i,j}}{\vartheta}\right)^k \right) \\ &\quad - \sum_{j=1}^m \left(\frac{R_{T_j}}{\vartheta}\right)^k. \end{aligned} \quad (12)$$

The maximum likelihood estimations for the parameters are defined by

$$\frac{1}{\hat{k}} + \frac{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} \log x_{i,j}}{\sum_{i=1}^m N_{T_j}} - \frac{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}} \log x_{i,j} + \sum_{j=1}^m R_{T_j}^{\hat{k}} \log R_{T_j}}{\sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}} + \sum_{j=1}^m R_{T_j}^{\hat{k}}} = 0 \quad (13)$$

and

$$\hat{\vartheta} = \left[\frac{\sum_{j=1}^m R_{T_j}^{\hat{k}} + \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} x_{i,j}^{\hat{k}}}{\sum_{j=1}^m N_{T_j}} \right]^{1/\hat{k}} \quad (14)$$

To apply the concept of minimum contrast estimation to this case we remark, that

$$\varrho(X_{i,j}, \theta) = - \left(\log k - \log \vartheta + (k-1)(\log X_{i,j} - \log \vartheta) - \left(\frac{X_{i,j}}{\vartheta} \right)^k \right)$$

is a contrast function (see example 1) and

$$\tilde{\varrho}(R_{T_j}) = - \left(\frac{R_{T_j}}{\vartheta} \right)^k$$

is a measurable function. Both $\varrho(x, \theta)$ and $\tilde{\varrho}(x, \theta)$ are continuous functions on $[0, \infty) \times \Theta$ with values in $(-\infty, +\infty)$ and twice continuously differentiable with respect to each θ_j , $j = 1, 2$, in Θ^0 .

To prove the other conditions of Theorems 1 and 2 the matrixes Σ and $\tilde{\Sigma}$ must be calculated. In the case of the Weibull distribution the Rao-Cramer conditions of regularity are fulfilled and it is easy to prove, that $\tilde{\Sigma} = \Sigma$. Σ is the Fisher information matrix of one observation and can be calculated in the following way:

$$\begin{aligned} \varrho_{1,1}(X_{1,1}, \theta) &= \frac{1}{k^2} \left(1 + \left(\frac{X_{i,j}}{\vartheta} \right)^k \log^2 \left(\frac{X_{i,j}}{\vartheta} \right)^k \right) \\ \varrho_{1,2}(X_{1,1}, \theta) &= \frac{1}{\vartheta} \left(1 - \left(\frac{X_{i,j}}{\vartheta} \right)^k \log \left(\frac{X_{i,j}}{\vartheta} \right)^k - \left(\frac{X_{i,j}}{\vartheta} \right)^k \right) \\ \varrho_{2,2}(X_{1,1}, \theta) &= -\frac{k}{\vartheta^2} + \frac{k(k+1)}{\vartheta^2} \left(\frac{X_{i,j}}{\vartheta} \right)^k. \end{aligned}$$

Using the equations $E\left(\frac{\partial \varrho(X_{1,1}, \theta)}{\partial \theta_k}\right) = 0$, $k = 1, 2$, after a substitution under the integral the information of one realization takes the form

$$\begin{aligned} \Sigma_{11} &= \frac{1}{k^2} \left(1 + \int_0^\infty u (\log u)^2 e^{-u} du \right) \\ \Sigma_{12} &= \frac{1}{\vartheta} \left(1 - \int_0^\infty u e^{-u} \log u du - \int_0^\infty u e^{-u} du \right) \\ \Sigma_{22} &= -\frac{k}{\vartheta^2} + \frac{k(k+1)}{\vartheta^2} \int_0^\infty u e^{-u} du. \end{aligned}$$

The integrals can be calculated using the gamma function and its derivations. Finally, we get

$$\Sigma = \begin{bmatrix} \frac{\pi^2/6 + (1-c)^2}{k^2} & -\frac{1-c}{\theta} \\ -\frac{1-c}{\theta} & \frac{k^2}{\theta^2} \end{bmatrix}$$

where c is Euler's constant $c \approx 0,5772$.

Now it is easy to see that the conditions of the Theorems 1 and 2 are fulfilled, i.e. $\hat{\theta}_n$ from (13) and (14) is weakly consistent and the distribution of

$$\sqrt{T_n}(\hat{\theta}_n - \theta)$$

converges weakly to a 2-dimensional normal distribution with expectation zero and covariance matrix $C(\theta) = \frac{1}{\mu}\Sigma^{-1}(\theta)$ where μ is the expectation of the Weibull distribution

$$\mu = \vartheta\Gamma\left(1 + \frac{1}{k}\right).$$

The fact that the maximum likelihood estimation is asymptotically normal distributed can be used to construct simultaneous confidence regions for the parameters of the Weibull distribution.

A level line of the two-dimensional Gaussian distribution can be determined such that inside of this line there is $(1 - \alpha)$ of the probability mass of the distribution. The equation of this level line is given by the quadratic form

$$(\hat{\theta} - \theta)^T \Sigma^{-1}(\hat{\theta} - \theta) < a^2 \quad (15)$$

with $\hat{\theta} = (\hat{k}, \hat{\vartheta})$, $\theta = (k, \vartheta)$, where $a^2 = -2 \log \alpha$ is the quantil of the χ^2 -distribution with 2 degrees of freedom and Σ is the Fisher information matrix. To calculate such a level line numerical methods must be used.

To compare the maximum likelihood estimator with the minimum contrast estimator considered in example 2 we investigate the scale model $f_\theta(x) = \theta f(\theta x)$, $\theta \in [a, b] \subseteq (0, \infty)$. Put $\varrho(x, \theta) = -\log \theta + x\theta$ and $\tilde{\varrho}(x, \theta) = \theta x$. Then

$$Q(\theta) = \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} (-\log \theta + x_{i,j}\theta) + \sum_{j=1}^m \theta R_{T_j}.$$

The minimum contrast estimation is then given by

$$\tilde{\theta} = \frac{\sum_{j=1}^m N_{T_j}}{\sum_{j=1}^m T_j}.$$

Suppose f is twice continuously differentiable and $\int_0^\infty t^2 f(t) dt < \infty$. We obtain from Theorem 2 that for $T_n \rightarrow \infty$ the distribution of the corresponding $\tilde{\theta}_n$ is asymptotically normal with expectation θ and variance $T_n \sigma^2 \theta^5$ where $\sigma^2 = \int_0^\infty (t - \frac{1}{\theta})^2 f(t) dt$. Now we specialize f and set

$$f(x) = \frac{2}{\sqrt{2\pi}} x^2 \exp\left(-\frac{x^2}{2}\right) \quad \text{for } x \geq 0$$

and $f(x) = 0$ for $x < 0$. Calculating σ^2 one easily derives that the distribution of $\sqrt{T_n}(\hat{\theta}_n - \theta)$ converges weakly to a normal distribution with mean zero and variance

$$C(\theta) = \theta^3 \cdot 2 \left(\sqrt{\pi/2} - 1 \right).$$

Otherwise the loglikelihood function is given by

$$\begin{aligned} \log L = & \sum_{j=1}^m \sum_{i=1}^{N_{T_j}} \left(\log \frac{2}{2\pi} + 3 \log \theta - 2 \log x_{i,j} - \frac{\theta^2 x_{i,j}^2}{2} \right) \\ & + \sum_{j=1}^m \log[1 - F(R_{T_j})], \end{aligned}$$

where $F(\cdot)$ is the distribution function of the Maxwell distribution. This distribution function can be calculated using normal distribution. But to found the maximum likelihood estimation observing renewal processes numerical methods have to be used.

To find asymptotic confidence intervals for the parameter θ the asymptotic properties of minimum contrast estimations can be used also. It is easy to see that the conditions of the Theorems 1 and 2 are fulfilled and the asymptotic variance of the maximum likelihood estimation is given by

$$C(\theta) = \theta^3 \frac{\sqrt{\pi/2}}{12}.$$

The maximum likelihood estimation gives a shorter confidence interval but is more complicated to calculate as the minimum contrast estimation from example 2.

References

- [1] Gut, A. : *Stopped Random Walks. Limit Theorems and Applications*. New York 1988
- [2] Ibragimov, I.A. and Has'minskii, R.Z. : *Statistical Estimation: Asymptotic Theory*. New York 1987
- [3] Liptser, R.S. and Shirayayev, A.N. : *Statistics of random processes II. Applications*. New York 1978
- [4] Mann, N.R., Schafer, R.E. and Singpurwalla, N.P. : *Methods for Statistical Analysis of Reliability and Life Data*. New York 1974
- [5] Pfanzagl, J. : *On the measurability and consistency of minimum contrast estimates*. *Metrika* **14**, 249-276 (1969)

[6] **Revesz, P.** : *Die Gesetze der großen Zahlen*. Budapest 1968

[7] **Serfling, R.** : *Approximation Theorems of Mathematical Statistics*. New York 1980

received: September 9, 1991

revised: January 9, 1992

Author:

PD Dr. W. Kahle
Technische Universität „Otto von Guericke“
Fakultät für Mathematik
Universitätsplatz 2
39106 Magdeburg
Germany

Prof. Dr. F. Liese
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

JENS MEYER

Epidemiemodelle und Iterierte Funktionssysteme¹

1 Einleitung

Die Ausbreitung ansteckender Krankheiten innerhalb einer Population wird klassischerweise durch Modelle beschrieben, die Differentialgleichungssysteme, stochastische Prozesse oder sogenannte Ketten-Binomial-Modelle verwenden. In unterschiedlichen Situationen kann man dadurch zu befriedigenden Resultaten gelangen. Dies ist jedoch nicht der Fall, wenn man die Krankheitszahlen über einen langen Zeitraum simulieren will, wobei ein Modell derartige Zeitreihen (vgl. [9]) beschreiben soll, wie sie in den Abb.1-4 gegeben sind. Eine Möglichkeit dazu bietet sich mit Iterierten Funktionensystemen. Dieser Artikel soll die grundlegenden Ideen dieses Ansatzes verdeutlichen. Bei der Betrachtung der Abbildungen fällt zudem auf, daß die „Formen“ der Kurven sich von Krankheit zu Krankheit unterscheiden. Dieser Aspekt soll bei der Modellierung mit berücksichtigt werden.

¹Ausgearbeitete Fassung eines Vortrags, gehalten auf dem Bremen-Rostocker-Statistik-Seminar, Rostock, 5.-7.März 1992

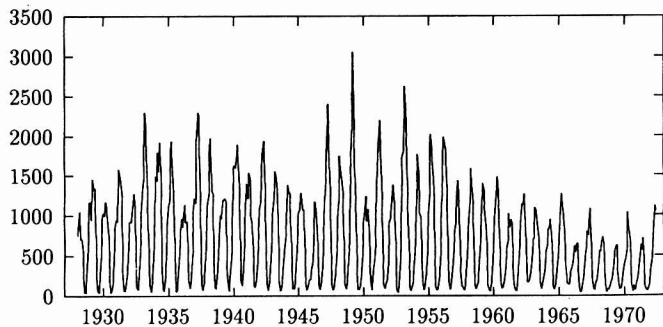


Abb.1: Windpockenfälle in New York 1928-1972

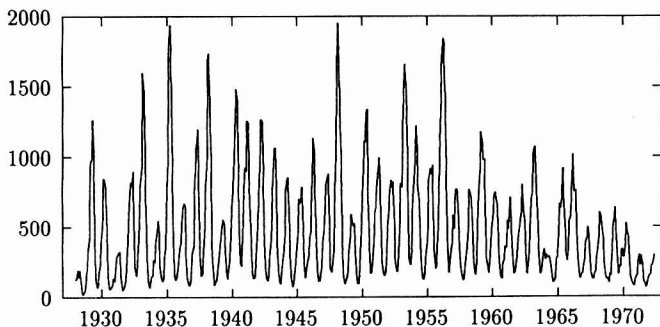


Abb.2: Mumpsfälle in New York 1928-1972

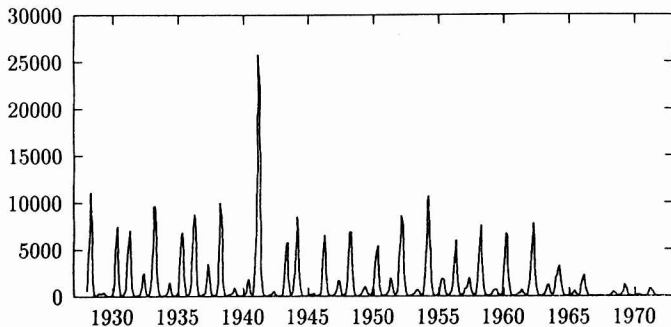


Abb.3: Masernfälle in New York 1928-1972

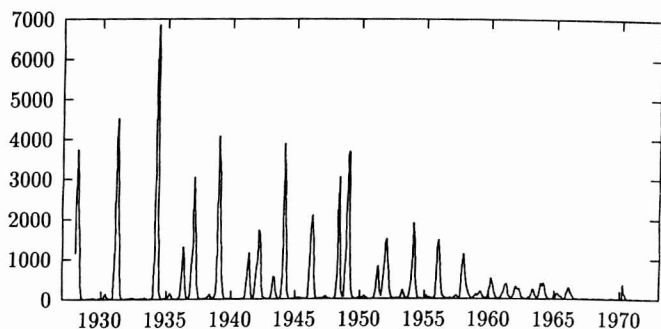


Abb.4: Masernfälle in Baltimore 1928–1972

2 Iterierte Funktionensysteme

Wir betrachten Abbildungen $T : \mathbb{R}_+^m \rightarrow \mathbb{R}_+^m$, die durch $x(n+1) = T(x(n))$ für alle $n \in \mathbb{N}$ gegeben sind. Die Abbildungen sollen genau eine kompakte Menge D besitzen, für welche $T(D) \subset D$ gilt (d.h., D ist bezüglich T positiv invariant). Weiterhin muß T im Folgenden immer eine Kontraktion sein, also

$$d(T(x), T(y)) \leq c \cdot d(x, y)$$

für eine Metrik d , für alle $x, y \in D$ und einen Kontraktionsfaktor $0 < c < 1$.

Banachscher Fixpunktsatz: Es seien (M, d) ein vollständiger metrischer Raum und $T : M \rightarrow M$ eine Kontraktion, dann gilt:

- a) T besitzt einen eindeutigen Fixpunkt $x^* \in M$ und
- b) $T^n(x(0)) \xrightarrow{n \rightarrow \infty} x^*$ für alle $x(0) \in M$, oder anders formuliert:

$$d(T^n(x(0)), x^*) \xrightarrow{n \rightarrow \infty} 0.$$

Man beachte, daß diese Aussagen bezüglich der Metrik gelten, für die T eine Kontraktion ist! Darüber hinaus wird die Bezeichnungsweise $T^0 = \text{id}$, $T^1 = T$, $T^{n+1} = T^n \circ T$ verwendet.

Wir betrachten nun N Abbildungen $T_i, i = 1, \dots, N$ mit den obigen Eigenschaften und die Iteration

$$\dots \circ T_1 \circ T_N \circ \dots \circ T_2 \circ T_1 \circ T_N \circ \dots \circ T_2 \circ T_1(x(0)).$$

Dabei seien die Fixpunkte der einzelnen T_i alle voneinander verschieden. Die Abbildung $T_N \circ T_{N-1} \circ \dots \circ T_2 \circ T_1$ besitzt einen eindeutigen Fixpunkt, da die Verknüpfung von Kontraktionen wiederum eine Kontraktion ist. Was passiert jedoch, wenn man alle Abbildungen

immer wieder einzeln, in einer festen Reihenfolge, nacheinander anwendet? Man erhält *keine* Konvergenz gegen einen Fixpunkt, sondern gegen N Häufungspunkte. Diese ergeben sich durch die Fixpunkte von

$$\begin{aligned} T_N \circ \cdots \circ T_2 \circ T_1(x(0)) \\ T_1 \circ T_N \circ \cdots \circ T_2(x(0)) \\ T_2 \circ T_1 \circ T_N \circ \cdots \circ T_3(x(0)) \\ \vdots \\ T_{N-1} \circ \cdots \circ T_1 \circ T_N(x(0)). \end{aligned}$$

Es erfolgt also eine Konvergenz gegen einen Orbit mit der Periode N .

Definition: Ein Orbit $\{x(0), T(x(0)) = x(1), T^2(x(0)) = T(x(1)) = x(2), \dots\}$ besitzt die *Periode* k , wenn $T^k(x(n)) = x(n)$ für das kleinste k mit dieser Eigenschaft gilt. Ein Punkt $x(0)$ heißt *schließlich periodisch*, wenn $T^m(x(0))$ für eine natürliche Zahl m periodisch ist.

Gegeben sind zwei verschiedene Startwerte $x(0)$ und $y(0)$. Wendet man auf diese beiden Startwerte dieselbe Folge von Abbildungen $T_{i_j}, i_j \in \{1, \dots, N\}$ an, so geht die Differenz der Bildpunkte $\cdots \circ T_{i_3} \circ T_{i_2} \circ T_{i_1}(x(0))$ und $\cdots \circ T_{i_3} \circ T_{i_2} \circ T_{i_1}(y(0))$ für $j \rightarrow \infty$ und $j \in \mathbb{N}$ gegen 0. Diese Eigenschaft wird als *Pfadstabilität* oder *schwache Ergodizität* bezeichnet. Insbesondere kehrt der Orbit auch nach einer Störung zu seinem ursprünglichen Verhalten zurück. Falls immer wieder dieselbe Abbildung angewendet wird, so nähern sich die beiden Orbits natürlich einem gemeinsamen Fixpunkt. Durchlaufen die i_j periodisch immer dieselbe Sequenz, so nähern sich die beiden Orbits einem periodischen Orbit. Ist die Folge der i_j jedoch willkürlich, so nähern sich die beiden Orbits einer Menge von Häufungspunkten, die oft ein Fraktal ist.

Es folgt nun der Fall, in dem wir N Kontraktionen T_i haben, wobei jeder Abbildung eine Wahrscheinlichkeit p_i zugeordnet wird. Die Anwendung der Transformationen verläuft folgendermaßen: Man wählt einen Anfangswert $x(0)$ und wählt dann stochastisch unabhängig entsprechend der Wahrscheinlichkeiten p_i eine der Abbildungen aus, so daß sich ein Bildpunkt $T_{i_1}(x(0))$ ergibt ($i_1 \in \{1, \dots, N\}$). Dann wird wieder eine Abbildung zufällig ausgewählt, so daß man einen nächsten Bildpunkt $T_{i_2}(T_{i_1}(x(0)))$ erhält. Diese Prozedur wird nun immer wieder durchgeführt, so daß sich schließlich eine oben angesprochene Menge \mathcal{A} von Häufungspunkten ergibt. Die Kontraktionen mit den entsprechenden Wahrscheinlichkeiten werden als *Iteriertes Funktionensystem mit Wahrscheinlichkeiten*, oder kurz als *IFS*, bezeichnet.

Dieselbe Menge \mathcal{A} von Häufungspunkten läßt sich auch noch ganz anders charakterisieren: (vgl. [6])

Satz von Hutchinson: Es seien (X, d) ein vollständiger metrischer Raum und $T_i : X \rightarrow X, i \in \{1, \dots, N\}$, Kontraktionen. Dann existiert genau eine invariante kompakte Menge

$A \neq \emptyset$, d.h.

$$A = \bigcup_{i=1}^N T_i(A).$$

Definition: Eine abgeschlossene Menge S heißt *Attraktor* für ein diskretes dynamisches System $T : D \rightarrow D$ mit $x(n+1) = T(x(n))$, wenn

- $\min\{d(T^n(x(0)), s); s \in S\} \xrightarrow{n \rightarrow \infty} 0$ für alle $x(0) \in D$ und
- $T^m(S) = S$ für alle $m \in \mathbb{N}$

gelten.

Insbesondere sind also auch Fixpunkte und periodische Orbits Attraktoren. IFS mit Wahrscheinlichkeiten besitzen jedoch meist Attraktoren, die aus unendlich vielen Punkten bestehen, welche *nicht* periodisch sind. Die Form des Attraktors ist durch die T_i bestimmt.

Es sei nun i_1, i_2, i_3, \dots eine zufällige Zahlenfolge, wobei i_j mit der Wahrscheinlichkeit p_k stochastisch unabhängig aus der Zahlenmenge $\{1, \dots, N\}$ für das Ereignis $i_j = k$ ausgewählt wird.

Was läßt sich über den Attraktor eines IFS mit Wahrscheinlichkeiten sagen?

- Es sei z^* ein Fixpunkt von einer der Abbildungen T_i . Dann gehören die Punkte $z^* = z(0), z(1) = T_{i_1}(z(0)), z(2) = T_{i_2}(z(1)), \dots$ zum Attraktor.
- Die Folge $z(0), z(1), z(2), \dots$ füllt den Attraktor fast sicher dicht aus. Solche Folgen (bzw. Orbits) heißen ergodisch. Die Ausnahme bilden periodische Folgen, jedoch ist die Wahrscheinlichkeit für ihr Auftreten gleich null. Ein ergodischer Orbit kommt jedem Punkt des Attraktors irgendwann beliebig nahe.
- Der Attraktor ist der topologische Abschluß aller periodischen Punkte. Er ist damit auch der Abschluß der Fixpunkte von allen endlichen Kollektionen von Kombinationen T_i .

Bemerkung: Der Beweis für diese Aussagen läßt sich relativ einfach durchführen, wenn man den *Adressenraum* des Attraktors verwendet. Dieses Konzept wird z.B. in [11] und [2] erläutert.

Meist werden nicht alle Teile des Attraktors gleich häufig „besucht“, sondern sehr ungleichmäßig. Dieser wichtigen Tatsache kann mit folgendem Konzept nachgegangen werden: Wenn $T : D \rightarrow D$ ein diskretes dynamisches System ist, so kann man für Teilmengen $A \subset D$ ein Maß

$$\mu(A) = \lim_{m \rightarrow \infty} \frac{1}{m} \cdot \#\{n : 1 \leq n \leq m \text{ und } T^n(x) \in A, x \in D\}$$

definieren, welches invariant ist, d.h., es gilt: $\mu(T^{-1}(A)) = \mu(A)$. Ein diskretes dynamisches System mit diesem invarianten Maß μ heißt ergodisch, wenn fast alle Orbits das Maß μ erzeugen. Fast alle Orbits eines IFS mit Wahrscheinlichkeiten sind ergodisch. Insbesondere wird bei ihnen das Maß μ durch die Wahrscheinlichkeiten bestimmt. Eine weitergehende Untersuchung dieses invarianten Maßes läßt sich mit Zuhilfenahme der Theorie der *Multifraktale* durchführen. Die geometrische Struktur von Fraktalen wird in [4] untersucht. Weitere Anwendungen von IFS auf biomathematische Zusammenhänge finden sich in [3] und [8].

3 Ein deterministisches Epidemiemodell

Das folgende $S - I - R - S$ -Modell beschreibt den zeitlichen Verlauf einer ansteckenden Krankheit, die von Mensch zu Mensch durch Tröpfcheninfektion übertragen wird. Die Population hat eine konstante Größe, und die durch die Krankheit erworbene Immunität ist nicht lebenslang. S bezeichnet den Anteil der ansteckbaren, I den Anteil der infektiösen und R den Anteil der immunen Personen. βSI beschreibt den Vorgang der Ansteckung, wobei davon ausgegangen wird, daß alle Personen hinsichtlich ihrer biologischen Eigenschaften gleich sind und sich gleichmäßig vermischen. β spiegelt das Maß an Infektiösität wider, $1/\gamma$ ist die mittlere Dauer der infektiösen Periode, $1/\delta$ ist die mittlere Dauer der immunen Phase und $1/\mu$ ist die mittlere Lebenserwartung, bzw. ist μ gleichzeitig die Geburtenrate. Dann wird dieses Modell durch folgendes Differentialgleichungssystem beschrieben:

$$\begin{aligned}\frac{dS(t)}{dt} &= \mu - \mu \cdot S(t) + \delta \cdot R(t) - \beta \cdot S(t) \cdot I(t) \\ \frac{dI(t)}{dt} &= \beta \cdot S(t) \cdot I(t) - \mu \cdot I(t) - \gamma \cdot I(t) \\ \frac{dR(t)}{dt} &= \gamma \cdot I(t) - \mu \cdot R(t) - \delta \cdot R(t).\end{aligned}$$

Eine Summation der drei Gleichungen ergibt $S(t) + I(t) + R(t) = 1$, so daß eine der Gleichungen überflüssig ist.

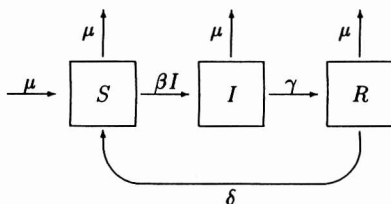


Abb.5: Flußdiagramm für eine Epidemie mit begrenzter Immunität.

Das Verhalten des Systems für $t \rightarrow \infty$ wird im wesentlichen durch einen Schwellenwert bestimmt: Entweder ergibt sich ein infektionsfreier Gleichgewichtspunkt (d.h. mit $I = 0$) oder ein endemisches Gleichgewicht (d.h. $I > 0$ für alle t). Das System bleibt immer beschränkt, und keine Komponente wird jemals kleiner als null. Bei „sinnvollen“ Anfangswerten ist das eine oder das andere Gleichgewicht, je nach Parameterkonstellation, somit asymptotisch stabil. Ein typisches Bild einer in die S - I -Ebene projizierten Trajektorie ist in Abb.6 zu sehen. Die Parameter in diesem und in folgenden Beispielen sind fiktiv und in der Einheit (Woche) $^{-1}$ angegeben. Damit erhält man eine Epidemie, deren infektiöse Periode 10Tage und die immune Phase \approx 1Jahr beträgt. Die Lebenserwartung eines Neugeborenen ist \approx 68.5Jahre. $\beta = 0.9$ hat nicht so eine einfache Bedeutung. Es gilt jedoch, je höher β ist, umso ansteckender ist die Krankheit. Dieses Epidemiemodell wird in [5] genauer untersucht. Eine allgemeine Darstellung der Seuchenausbreitung liefert [1].

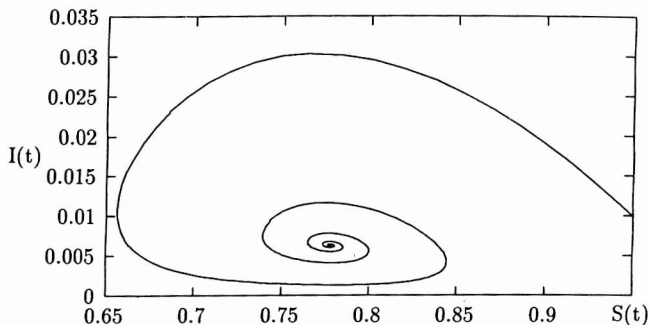


Abb.6: Projektion einer Trajektorie des SIRS-Modells in die S - I -Ebene. Die Parameter sind $\beta = 0.9$, $\gamma = 0.7$, $\mu = 0.00028$ und $\delta = 0.02$. Die Anfangswerte sind $S(0) = 0.95$, $I(0) = 0.01$ und $R(0) = 0.04$.

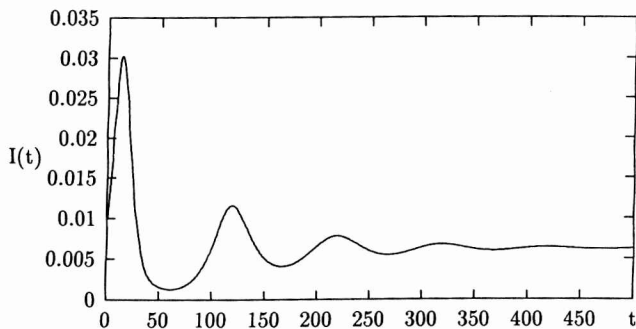


Abb.7: Verlauf von $I(t)$ gegen die Zeit (Parameter und Anfangswerte wie in Abb.6).

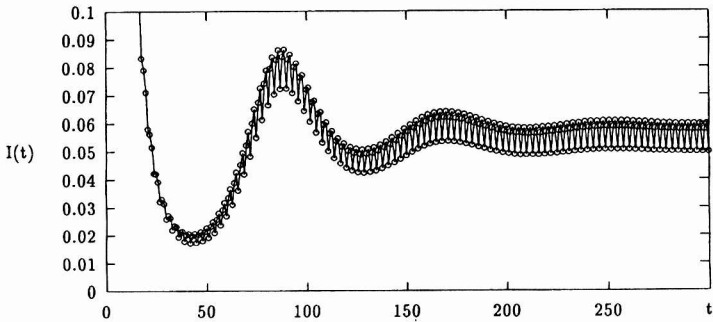


Abb.8: Die periodische Anwendung dreier Transformationen führt zu einem Orbit der Periode drei.

4 Epidemiemodelle mit Iterierten Funktionensystemen

Eine Diskretisierung des obigen Differentialgleichungssystems nach dem Euler-Verfahren erzeugt ein positives nichtlineares diskretes dynamisches System. Die Schrittweite h der Diskretisierung spiegelt den Vorhersagezeitraum wider. h läßt sich immer so klein wählen (man kann Schranken angeben), daß die ursprünglichen Eigenschaften des Differentialgleichungssystems (insbesondere die Gleichgewichtseigenschaften) unverändert bleiben. Es gelten weiterhin $S(n) + I(n) + R(n) = 1$ und $S(n), I(n), R(n) \geq 0$ für alle $n \in \mathbb{N}$. Wenn man also entsprechend vorgeht, so ergibt sich:

$$S(n+h) = S(n) + h(-\beta S(n)I(n) + \mu - \mu S(n) + \delta R(n))$$

$$I(n+h) = I(n) + h(\beta S(n)I(n) - \mu I(n) - \gamma I(n))$$

$$R(n+h) = R(n) + h(\gamma I(n) - \mu R(n) - \delta R(n)).$$

Wählt man dieselben Parameter wie in Abb.6 und setzt $h = 1$ Woche, so ergeben sich dieselben Kurven wie in den Abb.6 und 7.

Wir betrachten nun eine endliche Anzahl von Differentialgleichungssystemen von genau der oben beschriebenen Bauart, die sich nur hinsichtlich der Parameter unterscheiden. Jedes System wird diskretisiert und für sich als eine Abbildung $T : D \rightarrow D$ mit

$$\begin{pmatrix} S(n+1) \\ I(n+1) \\ R(n+1) \end{pmatrix} = T \begin{pmatrix} S(n) \\ I(n) \\ R(n) \end{pmatrix} \text{ angesehen. Dabei ist } D = \{(S, I, R) : S+I+R = 1 \text{ und } S, I, R \geq 0\}.$$

Diese Abbildungen werden in fester Reihenfolge oder zufällig immer wieder nacheinander angewendet. Damit ergibt sich ein IFS.

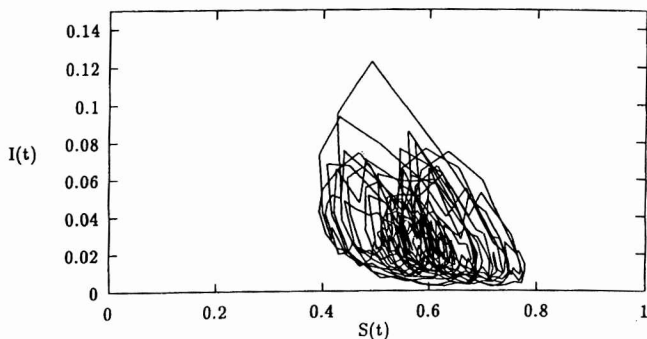


Abb.9: IFS mit vier Transformationen, die sich nur hinsichtlich des Parameters β unterscheiden. Die Wahrscheinlichkeit für die Anwendung sind für alle Transformationen gleich. Man sieht die Projektion eines Orbits in die S - I -Ebene.

Die Parameter werden bezüglich der Einheiten day^{-1} , week^{-1} , month^{-1} oder year^{-1} ausgedrückt. Dies sollte so geschehen, daß alle Parameter zwischen 0 und 1 liegen. Aus der Berechnung der Fixpunkte einer Abbildung ergibt sich ein Schwellenwert $\sigma = \frac{\beta}{\gamma + \mu}$, so daß folgende Aussagen gemacht werden können:

- 1) Für $\sigma \leq 1$ ist $\mathcal{G}_1 = (1, 0, 0)$ asymptotisch stabil auf D .
- 2) Für $\sigma > 1$ ist $\mathcal{G}_2 = \left(\frac{1}{\sigma}, \frac{(\mu + \delta)(\sigma - 1)}{\beta + \mu\sigma}, 1 - \frac{1}{\sigma} - \frac{(\mu + \delta)(\sigma - 1)}{\beta + \mu\sigma} \right)$ asymptotisch stabil auf $D \setminus \{(S, 0, R) : 0 \leq S, R \leq 1\}$.

Für ein IFS müssen die Abbildungen T_i Kontraktionen sein. Dies wird hier durch den „umgekehrten Banachschen Fixpunktsatz“ (vgl. [7]) abgesichert:

Satz: Es seien (X, d) ein kompakter metrischer Raum und T eine Abbildung mit $T(X) \subset X$, wobei die Iteration T^n von T für $n \rightarrow \infty$ für alle $x \in X$ gegen einen eindeutigen Fixpunkt x^* läuft. Dann existiert eine Metrik \tilde{d} , so daß T eine Kontraktion ist, d.h., es gilt $\tilde{d}(T(x), T(y)) \leq c \cdot \tilde{d}(x, y)$ mit $0 < c < 1$ für alle $x, y \in X$.

Die Metrik und die jeweiligen Kontraktionsfaktoren lassen sich für unsere konkreten Abbildungen explizit angeben. Aus ihnen lassen sich wichtige Eigenschaften des Attraktors ableiten. An dieser Stelle soll darauf jedoch nicht näher eingegangen werden.

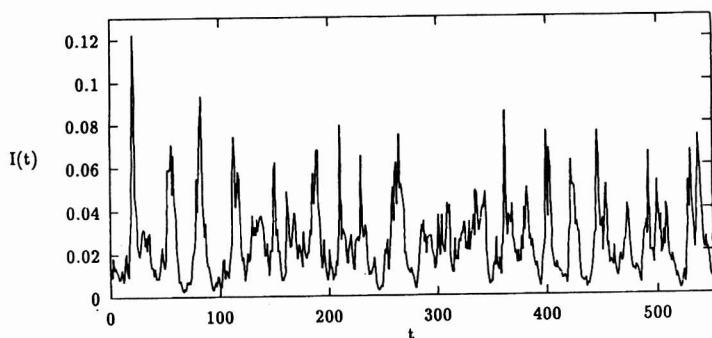


Abb.10: IFS mit denselben Parametern wie in Abb.9. Man sieht die Fluktuation der Infektiösen $I(t)$ mit der Zeit.

Es folgen nun einige Beispiele, die belegen sollen, daß diese Modellart Kurven der gewünschten Form erzeugen kann. Die jeweiligen Parameter sind in der Tabelle am Ende dieses Abschnitts zusammengestellt (h ist eine Woche). Sie sind in der Einheit (Woche) $^{-1}$ angegeben. In den Abbildungen, in denen $I(t)$ gegen t dargestellt ist, ist die Zeitachse in Monaten beschriftet. In Abb.8 werden drei Transformationen periodisch angewendet. Nachdem das System eingeschwungen ist, bewegt es sich schließlich nur noch auf einem Orbit mit der Periode drei. Abb.9 liegt ein IFS mit vier Transformationen zugrunde, die sich nur im Parameter β unterscheiden. Sie werden mit gleicher Wahrscheinlichkeit angewendet. Dabei besitzen die ersten drei Transformationen einen Schwellenwert σ , der größer als eins ist, während der Schwellenwert der vierten Transformation kleiner als eins ist. Abb.10 zeigt das entsprechende Verhalten des infektiösen Anteils $I(t)$ der Population. Die Abb.11 und 12 werden durch dieselbe Parameterkonstellation erzeugt, wobei jetzt jedoch die vierte Transformation mit $\sigma < 1$ mit der Wahrscheinlichkeit 0.55 statt 0.25 angewendet wird. Das System bewegt sich dementsprechend viel näher am epidemiefreien Zustand. Für Abb.13 wird wieder ein IFS mit vier Transformationen verwendet, die sich nun in allen vier Parametern unterscheiden. Insbesondere wird durch die vierte Transformation eine lebenslange Immunität modelliert ($\delta_4 = 0$). Verlängert man die Dauer der mittleren immunen Phase auf das Zehnfache und behält die anderen Parameter bei, so ergibt sich der Verlauf in Abb.14. Verständlicherweise verringert sich dadurch der Anteil infektiöser Personen, obwohl es weiterhin zu spontanen (etwas niedrigeren) Epidemieausbrüchen kommen kann. Die σ_i der Transformationen aus Abb.13 und Abb.14 unterscheiden sich jedoch nicht!

Die Entstehung der fraktalen Kurven läßt sich einfach als ein zufälliges „Umherspringen“

zwischen den jeweiligen Phasenräumen der Transformationen erklären. Dadurch ergeben sich die Möglichkeiten für eine Anpassung an reale Daten, Prognosen und Entwicklung von Impfstrategien. Dafür ist es nicht nur notwendig, den Anteil der infektiösen Personen zu verringern, sondern auch den der Ansteckbaren. Anschaulich bedeutet das, daß man z.B. das „Knäuel“ in Abb.11 durch eine aktive Immunisierung in die *linke* untere Ecke des Diagramms bringen muß.

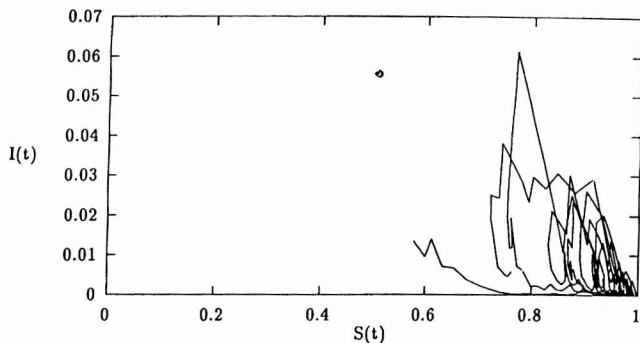


Abb.11: IFS mit denselben Parametern wie in Abb.9. Die Wahrscheinlichkeit für die Anwendung der Transformation, deren Schwellenwert σ kleiner als eins ist, liegt jedoch bei 0.55.

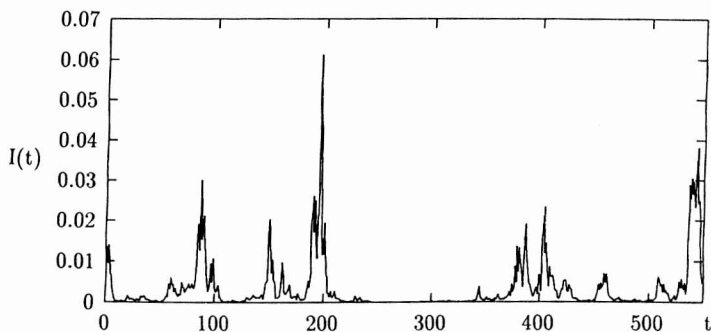


Abb.12: Verlauf von $I(t)$ gegen die Zeit für das IFS aus Abb.11.

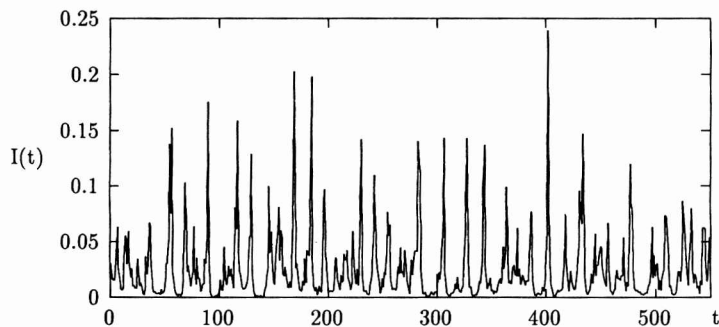


Abb.13: Die Transformationen dieses IFS unterscheiden sich in allen Parametern.

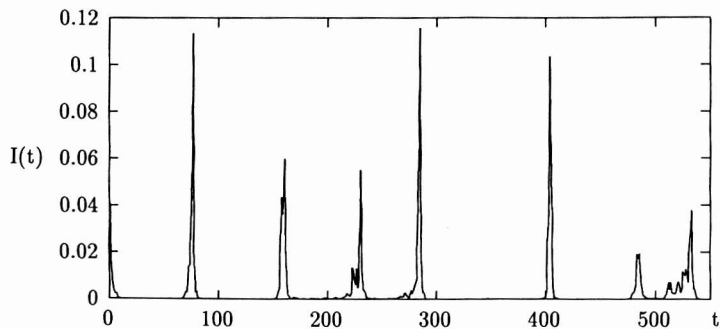


Abb.14: Bis auf δ sind die Parameter wie in Abb.13 gewählt. Die immune Phase ist hier zehnmal so lang.

Parameterwerte der Transformationen für die Beispiele in den Abbildungen

	Abb.8	Abb.9 und 10	Abb.11 und 12	Abb.13	Abb.14
p_1	—	0.25	0.15	0.2	0.2
p_2	—	0.25	0.15	0.4	0.4
p_3	—	0.25	0.15	0.2	0.2
p_4	—	0.25	0.55	0.2	0.2
β_1	0.9	0.9	0.9	0.95	0.95
β_2	0.6	0.7	0.7	0.7	0.7
β_3	0.1	0.4	0.4	0.95	0.95
β_4	—	0.1	0.1	0.75	0.75
γ_1	0.5	0.3	0.3	0.25	0.25
γ_2	0.5	0.3	0.3	0.8	0.8
γ_3	0.5	0.3	0.3	0.3	0.3
γ_4	—	0.3	0.3	0.5	0.5
μ_1	0.00028	0.00028	0.00028	0.00032	0.00032
μ_2	0.00028	0.00028	0.00028	0.00028	0.00028
μ_3	0.00028	0.00028	0.00028	0.00025	0.00025
μ_4	—	0.00028	0.00028	0.00029	0.00029
δ_1	0.02	0.02	0.02	0.025	0.0025
δ_2	0.02	0.02	0.02	0.09	0.009
δ_3	0.02	0.02	0.02	0.075	0.0075
δ_4	—	0.02	0.02	0.00	0.00

5 Schlußbemerkungen

Die biologische Begründung für die Iteration verschiedener Abbildungen liegt in der relativ einfachen Darstellbarkeit der zeitlichen Schwankung von Parametern. Zeitlich periodische Schwankungen lassen sich z.B. mit jahreszeitlichen Änderungen der Umgebung erklären, zufällige Schwankungen durch Meßungenauigkeiten und eine Varianz der Parameter an sich. Man kann sich diesen Ansatz als eine diskrete Formulierung eines nicht-autonomen Differentialgleichungssystems vorstellen. Zusätzlich ergibt sich hierbei die Möglichkeit der Modellierung einer zufälligen Parameterfluktuation, welche im Gegensatz zu stochastischen Differentialgleichungen (d.h. „deterministischer Anteil plus Rauschen“) mit endlich vielen Zufallsausprägungen auskommt.

Die Anpassung eines IFS mit Wahrscheinlichkeiten an eine epidemiologische Zeitreihe soll an dieser Stelle nicht näher erläutert werden. Im wesentlichen wird die reale Zeitreihe in einem höherdimensionalen Raum eingebettet und dadurch die Trajektorie des zugrundeliegenden dynamischen Systems rekonstruiert (vgl. [13]). Es ergibt sich ein Attraktor, welcher mit

dem Attraktor des IFS verglichen wird. Das kann mit statistischen Methoden und Hilfsmitteln aus der fraktalen Geometrie (z.B. Bestimmung von Hausdorff- und Informationsdimension, Lyapunov-Exponenten, Poincaré- und First-Return-Abbildungen) ausgeführt werden. Eine Auswahl dieser Verfahren wird z.B. in [12] und [10] vorgestellt. Das Ziel besteht in der Konstruktion einer bijektiven (vielleicht affin linearen) Transformation zwischen den Einbettungen der wirklichen und der modellierten Zeitreihe. Darüber hinaus lassen sich jeweils die stationären Verteilungen miteinander vergleichen.

Literaturverzeichnis

- [1] **Bailey, N.T.J.** : *The Mathematical Theory of Infectious Diseases*. London 1975
- [2] **Barnsley, M.** : *Fractals Everywhere*. San Diego 1988
- [3] **Caswell, H.** : *Matrix Population Models*. Sunderland (Mass.) 1989
- [4] **Falconer, K.** : *Fractal Geometry*. Chichester 1990
- [5] **Hethcote, H.** : *Qualitative analyses of communicable disease models*. *Math. Biosci.* **28**, 335–356 (1976)
- [6] **Hutchinson, J.** : *Fractals and self-similarity*. *Indiana Univ. Math. J.* **30**, 713–747 (1981)
- [7] **Janos, L.** : *A converse of Banach's contraction theorem*. *Proc. Amer. Math. Soc.* **18**, 287–289 (1967)
- [8] **Krause, U.** : *Über positive diskrete dynamische Systeme*. *Rostock. Math. Kolloq.* **44**, 29–38 (1991)
- [9] **London W.P. and Yorke, J.A.** : *Recurrent outbreaks of measles, chickenpox and mumps*. *Amer. J. of Epidem.* **98**, 453–482 (1973)
- [10] **Olsen, L.F. and Degn, H.** : *Chaos in biological time series*. *Quart. Rev. of Biophysics* **18**, 165–225 (1985)
- [11] **Peitgen, H.O., Jürgens, H. and Saupe, D.** : *Chaos and Fractals*. New York 1992
- [12] **Sugihara, G., Grenfell, B. and May, R.M.** : *Distinguishing error from chaos in ecological time series*. *Phil. Trans. Roy Soc. London B* **330**, 235–251 (1990)

- [13] **Takens, F.** : *Detecting strange attractors in turbulence.* Lect. Notes in Math. **898**, 366–381 (1981)

received: 28. September 1993

Author:

Dipl.-Math. J. Meyer
Universität Bremen
Fachbereich Mathematik und Informatik
Postfach 33 04 40
28334 Bremen
Germany

BURKHARD LENZE

Design and realization of a family of optimal windows

ABSTRACT. In this paper, we discuss the design and explicit realization of a family of optimal windows. The variational concept with respect to which the constructed windows are optimal is motivated by the observation that variational conditions on the derivatives of the impulse response of the window in the time-domain correspond to energy and/or amplitude moment conditions on the system function in the frequency-domain, and, minimizing energy and/or amplitude moments of the system function is a standard strategy in window design. Moreover, we compare the new windows with a number of well-known explicitly representable windows like Bartlett, Tukey, Blackman, Parzen, and Papoulis window in order to show that the new ones are at least as good as the standard ones, and, with respect to some performance criteria they are even better.

KEY WORDS. Signal modeling, optimal windows, explicit realization, amplitude and energy moments, asymptotic attenuation, perfect B-splines.

1 Introduction

As usual (cf. [12], p. 236), we say that the functions

$$w_\tau(t) \longleftrightarrow W_\tau(\omega) \quad (1)$$

with

$$W_\tau(\omega) := \int_{-\infty}^{\infty} w_\tau(t) e^{-j\omega t} dt \quad (2)$$

form a window pair if they are real and even and satisfy the following conditions:

(1) w_τ is normalized at the origin by

$$w_\tau(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_\tau(\omega) d\omega = 1. \quad (3)$$

(2) w_τ is time-limited on $[-\tau, \tau]$, i.e.,

$$w_\tau(t) = 0 \quad \text{for } |t| > \tau . \quad (4)$$

(3) W_τ is of "short duration" in some sense.

While the first two conditions are mathematically precise, condition (3) is somehow vague. Intuitively, the "best short duration" would be an arbitrary small interval $[-\omega_0, \omega_0]$ with

$$W_\tau(\omega) = 0 \quad \text{for } |\omega| > \omega_0 , \quad (5)$$

however, this is impossible, because w_τ cannot be time- and band-limited, simultaneously. Therefore, a meaningful measure for something like short duration could be the existence of a sufficiently large $n \in \mathbb{N}$ such that W_τ satisfies the so-called *asymptotic attenuation* condition (cf. [12], p. 237)

$$W_\tau(\omega) = O(\omega^{-n}) \quad \text{for } |\omega| \rightarrow \infty . \quad (6)$$

Roughly speaking, the measure of duration of W_τ is the order of decrease of $W_\tau(\omega)$ for $|\omega| \rightarrow \infty$. Similar measures can be obtained by requiring the existence of the n -th *absolute amplitude moment* of W_τ ,

$$m_n(w_\tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega|^n W_\tau(\omega) d\omega < \infty , \quad (7)$$

or the existence of the $2n$ -th *energy moment* of W_τ ,

$$M_{2n}(w_\tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^{2n} W_\tau^2(\omega) d\omega < \infty . \quad (8)$$

In any case, each of these conditions not only imply a kind of short duration of W_τ in the sense of sufficiently fast decrease for $|\omega| \rightarrow \infty$, but also implicitly a smoothness assumption on w_τ . For example, the fact that $W_\tau(\omega)$ and $\omega^n W_\tau(\omega)$ (comp. (7)) are Lebesgue integrable over $(-\infty, \infty)$ immediately implies that w_τ is n -times continuously differentiable on $(-\infty, \infty)$ and all derivatives of f may be calculated by

$$w_\tau^{(k)}(t) = \frac{j^k}{2\pi} \int_{-\infty}^{\infty} \omega^k W_\tau(\omega) e^{j\omega t} d\omega , \quad 0 \leq k \leq n , \quad (9)$$

(cf. for example [3], p. 197, Proposition 5.1.17 with f^\wedge instead of f). Summing up, we see that a sufficiently short duration of W_τ in the above sense is only possible to obtain by using functions w_τ which are sufficiently smooth on the whole real line.

Therefore, we now directly start with a function w_τ which is n -times continuously differentiable on $(-\infty, \infty)$. As shown above, (7) then implies (9), respectively,

$$w_\tau^{(k)}(t) \longleftrightarrow (j\omega)^k W_\tau(\omega), \quad 0 \leq k \leq n. \quad (10)$$

Finally, applying Parseval's equation to (10) we obtain

$$\int_{-\tau}^{\tau} \left(w_\tau^{(k)}(t) \right)^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^{2k} W_\tau^2(\omega) d\omega, \quad 0 \leq k \leq n, \quad (11)$$

(comp. (8)). In conclusion, equations (10) and (11) show that there is an intimate connection between the order of magnitude of the derivatives of w_τ and the order of magnitude of amplitude and energy moments of W_τ .

Now, taking a look at the literature, various sophisticated strategies have been developed in order to minimize (11) resp. (8) and/or (9) resp. (7) (for example, compare the textbooks [11,12,15] and the (survey) papers [1,4,10] on optimal windows). In this paper, we focus our attention on a kind of simultaneous uniform minimization of all integrals under consideration by minimizing the maximum absolute value of the highest possible derivative of w_τ which is – at least weakly – available. More precisely, our optimization problem may be formulated as follows.

Optimal Window Problem.

For fixed $n \in \mathbb{N}$ and $\tau > 0$ find a window pair $w_{\tau,n}(t) \leftrightarrow W_{\tau,n}(\omega)$ such that the derivatives $w_{\tau,n}^{(k)}$, $0 \leq k \leq n-1$, exist and are absolutely continuous on \mathbb{R} and that the essential supremum norm of $w_{\tau,n}^{(n)}$,

$$\|w_{\tau,n}^{(n)}\|_\infty^{\mathbf{R}} := \text{ess sup} \left\{ \left| w_{\tau,n}^{(n)}(t) \right| : -\infty < t < \infty \right\}, \quad (12)$$

is minimal under all functions w_τ satisfying the stated smoothness conditions.

Note that by means of Goldberg's inequalities (cf. [14], pp. 17ff., for example) the control of $\|w_\tau^{(n)}\|_\infty^{\mathbf{R}}$ together with $\|w_\tau\|_\infty^{\mathbf{R}} = w_\tau(0) = 1$ (i.g.) guarantees a kind of simultaneous control of all intermediate derivatives $\|w_\tau^{(k)}\|_\infty^{\mathbf{R}}$, $0 < k < n$. Therefore, the formulation of the Optimal Window Problem makes sense in view of a kind of simultaneous uniform minimization of all energy and amplitude moments which are available.

In Section 2, we will give the general solution of the above problem for arbitrary $n \in \mathbb{N}$. In Section 3, we will discuss the cases $n = 1, 2, 3$ which should be of primary interest in practice, in some detail. Moreover, we will show how these windows are connected with well-known extensively used windows like Bartlett, Tukey, Blackman, Parzen, and Papoulis window. This discussion will be continued in Section 4 where we compare all windows under consideration with respect to a number of optimality and performance criteria. The paper will end with some main conclusions in Section 5.

2 General Solution

In order to give the general solution of the Optimal Window Problem of Section 1 we need a few introductory definitions.

Let $n \in \mathbb{N}$ be given arbitrarily. With

$$(t)_+^n := \begin{cases} t^n & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \quad (13)$$

the so-called n -th truncated power and

$$t_k^{(n+1)} := -\cos \frac{k\pi}{n+1}, \quad 0 \leq k \leq n+1, \quad (14)$$

the extremal points of the $(n+1)$ -degree Čebyšev polynomial of first kind, we define the perfect B-spline b_n of degree n by

$$b_n(t) := (n+1) \sum_{k=0}^{n+1} \left\{ \prod_{\substack{r=0 \\ r \neq k}}^{n+1} (t_k^{(n+1)} - t_r^{(n+1)}) \right\}^{-1} (t_k^{(n+1)} - t)_+^n, \quad t \in \mathbb{R}, \quad (15)$$

(cf. [13]). It is well-known (see again [13] or [14], for example) that:

(1) b_n is $(n-1)$ -times continuously differentiable on \mathbb{R} ,

(2) $b_n^{(n-1)}$ is absolutely continuous on \mathbb{R} ,

(3) $b_n(t) \begin{cases} > 0 & \text{for } |t| < 1 \\ = 0 & \text{for } |t| \geq 1 \end{cases}$, (16)

(4) $\int_{-\infty}^{\infty} b_n(t) dt = \int_{-1}^1 b_n(t) dt = 1$, (17)

(5) $\int_{-1}^1 b_n(t) e^{-j\omega t} dt = \frac{2^{n-1} n!}{(j\omega)^{n+1}} \left(e^{j\omega} + 2 \sum_{k=1}^n (-1)^k e^{-j\omega t_k^{(n+1)}} + (-1)^{n+1} e^{-j\omega} \right)$, $\omega \neq 0$. (18)

In [9], we proved that the perfect B-splines solve a variational minimum problem which is intimately connected with the Optimal Window Problem of Section 1.

Theorem 1 *Let $n \in \mathbb{N}$ be given arbitrarily and K_n the set of functions $f: \mathbb{R} \rightarrow \mathbb{R}$ which are time-limited on $[-1, 1]$, which have absolutely continuous derivatives $f^{(k)}$, $0 \leq k \leq n-1$, on \mathbb{R} , and which are normalized at the origin by $f(0) = 1$. Then the normalized perfect B-spline \bar{b}_n ,*

$$\bar{b}_n(t) := (b_n(0))^{-1} b_n(t), \quad t \in \mathbb{R}, \quad (19)$$

has minimal n -th order derivative with respect to the essential supremum norm under all functions in K_n , i.e.,

$$\|\tilde{b}_n^{(n)}\|_{\infty}^{\mathbf{R}} \leq \|f^{(n)}\|_{\infty}^{\mathbf{R}} \quad (20)$$

for all $f \in K_n$. Moreover, \tilde{b}_n is the unique solution.

Proof. Compare [9], p. 302, Theorem 3.1 and, moreover, [2,5,6,7,8,13] for the basic ideas. ■

From Theorem 1, we easily deduce the solution of our Optimal Window Problem by means of a simple change of scale.

Corollary 2 *Let $n \in \mathbb{N}$ and $\tau > 0$ be given arbitrarily. Then the unique solution of the Optimal Window Problem in Section 1 is given by the pair*

$$w_{\tau,n}(t) := \tilde{b}_n\left(\frac{t}{\tau}\right) = (b_n(0))^{-1} b_n\left(\frac{t}{\tau}\right), \quad (21)$$

$$W_{\tau,n}(\omega) := \int_{-\infty}^{\infty} w_{\tau,n}(t) e^{-j\omega t} dt = \frac{1}{b_n(0)} \int_{-\tau}^{\tau} b_n\left(\frac{t}{\tau}\right) e^{-j\omega t} dt. \quad (22)$$

After this general solution, we are now interested in a more detailed analysis of those cases which should be important for practical use ($n = 1, 2, 3$) and in their explicit realization.

3 The first three special cases

The case $n = 1$: In this case, (14) and (15) reduce to

$$t_0^{(2)} = -1, \quad t_1^{(2)} = 0, \quad t_2^{(2)} = 1, \quad (23)$$

$$b_1(t) = (-1 - t)_+^1 - 2(-t)_+^1 + (1 - t)_+^1. \quad (24)$$

Explicitly, in this special situation b_1 is nothing else but the well-known hat function, precisely,

$$b_1(t) = \begin{cases} 1 - |t| & \text{for } |t| \leq 1 \\ 0 & \text{for } |t| > 1 \end{cases}. \quad (25)$$

Therefore, the optimal window pair ((21), (22)) is the Bartlett window with the Fejér kernel as system function,

$$w_{\tau,1}(t) = b_1\left(\frac{t}{\tau}\right), \quad (26)$$

$$W_{\tau,1}(\omega) = \frac{4 \sin^2\left(\frac{\tau\omega}{2}\right)}{\tau\omega^2}. \quad (27)$$

This result shows that the Bartlett window not only overcomes Gibbs' phenomenon but also is optimal with respect to our optimality condition. Perhaps, this gives an additional hint why the Bartlett window is used in practice in such an extensive way (comp. also the remarks of Papoulis [12], p. 237, concerning the Bartlett window and other windows with good overall properties but no (known) optimal behaviour).

The case $n = 2$:

In this case, (14) and (15) yield

$$t_0^{(3)} = -1, \quad t_1^{(3)} = -\frac{1}{2}, \quad t_2^{(3)} = \frac{1}{2}, \quad t_3^{(3)} = 1, \quad (28)$$

$$b_2(t) = -2(-1-t)_+^2 + 4\left(-\frac{1}{2}-t\right)_+^2 - 4\left(\frac{1}{2}-t\right)_+^2 + 2(1-t)_+^2. \quad (29)$$

Therefore, using (18) the optimal window pair ((21), (22)) is given by

$$w_{\tau,2}(t) = b_2\left(\frac{t}{\tau}\right), \quad (30)$$

$$W_{\tau,2}(\omega) = \frac{32(\sin \frac{\omega\tau}{2})(\sin^2 \frac{\omega\tau}{4})}{\tau^2\omega^3}. \quad (31)$$

Obviously, this optimal window precisely lies between Bartlett and Parzen window and is equivalent to Tukey and Blackman window in view of its asymptotic attenuation (for details, compare [12], pp. 237ff., and [4,10]; see also Section 4). Moreover, it may be easily realized as a cascade connection of an ideal low-pass and a Bartlett window (neglecting scaling).

The case $n = 3$:

In this case, (14) and (15) imply

$$t_0^{(4)} = -1, \quad t_1^{(4)} = -\frac{1}{\sqrt{2}}, \quad t_2^{(4)} = 0, \quad t_3^{(4)} = \frac{1}{\sqrt{2}}, \quad t_4^{(4)} = 1, \quad (32)$$

$$b_3(t) = 4(-1-t)_+^3 - 8\left(-\frac{1}{\sqrt{2}}-t\right)_+^3 + 8(-t)_+^3 - 8\left(\frac{1}{\sqrt{2}}-t\right)_+^3 + 4(1-t)_+^3. \quad (33)$$

Here, again with the help of (18), the optimal window pair ((21), (22)) explicitly reads as follows:

$$w_{\tau,3}(t) = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right)b_3\left(\frac{t}{\tau}\right), \quad (34)$$

$$W_{\tau,3}(\omega) = \frac{48(1 + \frac{1}{\sqrt{2}})(2\sin^2(\frac{\omega\tau}{2\sqrt{2}}) - \sin^2(\frac{\omega\tau}{2}))}{\tau^3\omega^4}. \quad (35)$$

By means of (35), we see that our optimal window is asymptotically equivalent to Parzen and Papoulis window (comp. again [12], pp. 237ff., and [4]; see also Section 4). Moreover, it may be realized as a parallel connection of two Bartlett windows followed by two ideal integrators (again, neglecting appropriate scaling).

4 Some comparisons

As in [12], for

$$w_\tau(t) \longleftrightarrow W_\tau(\omega) \quad (36)$$

a given window pair we will consider the following quantities, as far as they exist:

Energy:

$$E(w_\tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} W_\tau^2(\omega) d\omega = \int_{-\tau}^{\tau} w_\tau^2(t) dt, \quad (37)$$

Energy moment:

$$M_2(w_\tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 W_\tau^2(\omega) d\omega = \int_{-\tau}^{\tau} (w'_\tau(t))^2 dt, \quad (38)$$

Amplitude moment:

$$m_2(w_\tau) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 W_\tau(\omega) d\omega = -w''_\tau(0), \quad (39)$$

Relative energy moment:

$$r(w_\tau) := \frac{M_2(w_\tau)}{E(w_\tau)}. \quad (40)$$

In this terminology, the goodness of a given window pair is measured by the absolute values of the energy moment, amplitude moment, and relative energy moment. The window is the better the smaller these values are. Obviously, these measures depend heavily on the behaviour of the frequency function W_τ for large $|\omega|$. On the other hand, it is sometimes of even greater interest to know how W_τ behaves for small values of ω . This leads to so-called sidelobe-energy to mainlobe-energy resp. sidelobe-peak to mainlobe-peak resp. first null-location criteria (cf. [1,4,10] for a detailed discussion and motivation of these performance criteria). We will not give a detailed analysis with respect to these conditions, here, but will only take a look at some plots of our frequency functions $W_{\tau,n}$ in order to get at least some

visual idea of their behaviour in connection with these performance criteria. As usual in this context, we will not plot $W_{\tau,n}$, directly, but

$$W_{\tau,n}^{(dB)}(2\pi\omega) := 20 \log_{10} \left(\frac{|W_{\tau,n}(2\pi\omega)|}{|W_{\tau,n}(0)|} \right), \quad \omega \in \mathbf{R}, \quad (41)$$

for sake of a higher resolution and to make our plots immediately comparable to those in [4] and [10]. Finally, let us note that since our window $w_{\tau,1}(t) \leftrightarrow W_{\tau,1}(\omega)$ is equal to the Bartlett window which has already been considered in literature, extensively, we can restrict ourselves to the cases $n = 2$ and $n = 3$.

The case $n = 2$:

We compare our window $w_{\tau,2}(t) \leftrightarrow W_{\tau,2}(\omega)$ (see (30) and (31)) with the Tukey window (also called Hanning or Hann window),

$$u_{\tau}(t) := \begin{cases} \frac{1}{2} \left(1 + \cos \pi \frac{t}{\tau} \right) & \text{for } |t| \leq \tau \\ 0 & \text{for } |t| > \tau \end{cases} \quad (42)$$

$$U_{\tau}(\omega) = \frac{\pi^2 \sin \tau \omega}{\omega(\pi^2 - \tau^2 \omega^2)}, \quad (43)$$

and the Blackman window,

$$k_{\tau}(t) := \begin{cases} \left(0.42 + 0.5 \cos \pi \frac{t}{\tau} + 0.08 \cos 2\pi \frac{t}{\tau} \right) & \text{for } |t| \leq \tau \\ 0 & \text{for } |t| > \tau \end{cases}, \quad (44)$$

$$K_{\tau}(\omega) = \frac{2\pi^2(0.18\tau^2\omega^2 - 1.68\pi^2) \sin \tau \omega}{\omega(\pi^2 - \tau^2\omega^2)(\tau^2\omega^2 - 4\pi^2)}. \quad (45)$$

We choose these two windows for comparison since Tukey and Blackman window have the same asymptotic attenuation as our window, namely,

$$U_{\tau}(\omega) = O(\omega^{-3}), \quad K_{\tau}(\omega) = O(\omega^{-3}) \quad (\omega \rightarrow \pm\infty). \quad (46)$$

Now, for the three windows under consideration we obtain the following, where from now on the sign \simeq means that the constants appearing in the expressions have been rounded with respect to the last printed digit (see also [12], p. 237, and [4,10]):

$$E(w_{\tau,2}) = \frac{23\tau}{30}, \quad M_2(w_{\tau,2}) = \frac{8}{3\tau}, \quad m_2(w_{\tau,2}) = \frac{4}{\tau^2}, \quad r(w_{\tau,2}) = \frac{80}{23\tau^2}, \quad (47)$$

$$E(u_{\tau}) = \frac{3\tau}{4}, \quad M_2(u_{\tau}) = \frac{\pi^2}{4\tau}, \quad m_2(u_{\tau}) = \frac{\pi^2}{2\tau^2}, \quad r(u_{\tau}) = \frac{\pi^2}{3\tau^2}, \quad (48)$$

$$E(k_{\tau}) = 0.6092\tau, \quad M_2(k_{\tau}) = \frac{0.2756\pi^2}{\tau}, \quad m_2(k_{\tau}) = \frac{0.82\pi^2}{\tau^2}, \quad r(k_{\tau}) \simeq \frac{4.465}{\tau^2}. \quad (49)$$

As is easily seen, in view of the above criteria our window is better than the Blackman window; it is even better than the Tukey window with respect to m_2 , but a bit less as good in terms of M_2 and r (see also Figure 2, and Figures 1 and 2 in [10] for further qualitative comparison).

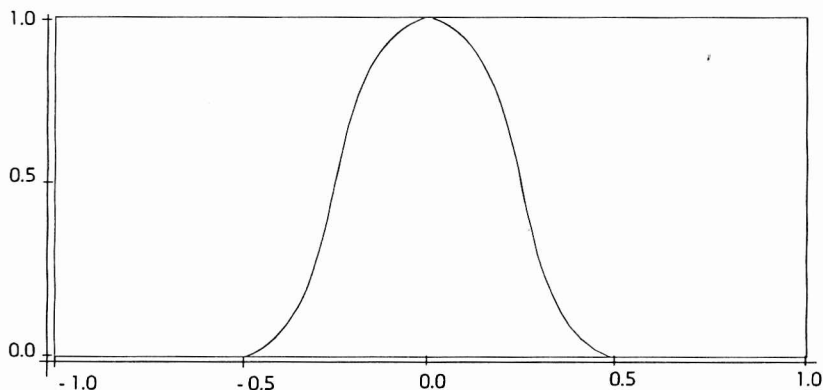


Figure 1. Plot of $w_{\tau,2}(t)$ for $\tau = \frac{1}{2}$ and $t \in [-1, 1]$.

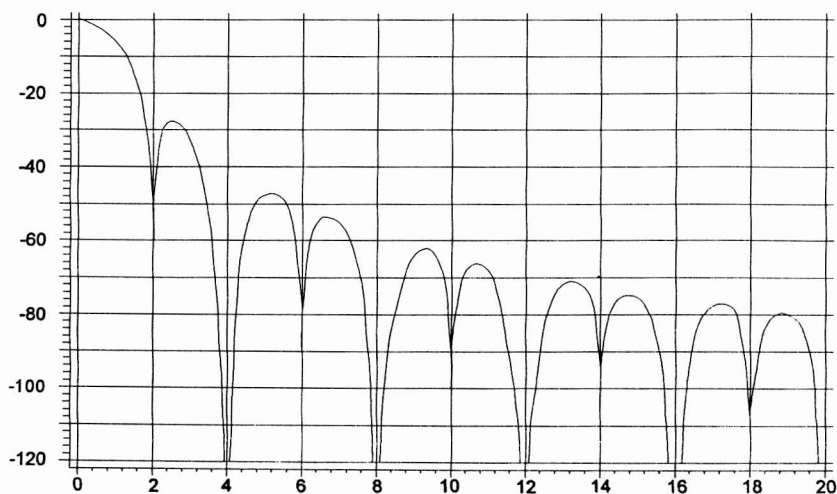


Figure 2. Plot of $W_{\tau,2}^{(dB)}(2\pi\omega)$ for $\tau = \frac{1}{2}$ and $\omega \in [0, 20]$.

The case $n = 3$:

We compare our window $w_{\tau,3}(t) \leftrightarrow W_{\tau,3}(\omega)$ (see (34) and (35)) with the Parzen window,

$$p_{\tau}(t) := 2 \left(-1 - \frac{t}{\tau}\right)_+^3 - 8 \left(-\frac{1}{2} - \frac{t}{\tau}\right)_+^3 + 12 \left(-\frac{t}{\tau}\right)_+^3 - 8 \left(\frac{1}{2} - \frac{t}{\tau}\right)_+^3 + 2 \left(1 - \frac{t}{\tau}\right)_+^3, \quad (50)$$

$$P_{\tau}(\omega) = \frac{3\tau}{4} \left(\frac{\sin \frac{\tau\omega}{4}}{\frac{\tau\omega}{4}}\right)^4, \quad (51)$$

and the Papoulis window (cf. [12], p. 239, and [4]),

$$a_{\tau}(t) := \begin{cases} \frac{1}{\pi} \left| \sin \pi \frac{t}{\tau} \right| + \left(1 - \frac{|t|}{\tau}\right) \cos \pi \frac{t}{\tau} & \text{for } |t| \leq \tau \\ 0 & \text{for } |t| > \tau \end{cases}, \quad (52)$$

$$A_{\tau}(\omega) = \frac{4\tau\pi^2(1 + \cos \tau\omega)}{(\pi^2 - \tau^2\omega^2)^2}. \quad (53)$$

Again, we choose these two windows for comparison since both they have the same asymptotic attenuation as our window, namely,

$$P_{\tau}(\omega) = O(\omega^{-4}), \quad A_{\tau}(\omega) = O(\omega^{-4}) \quad (\omega \rightarrow \pm\infty). \quad (54)$$

For these windows we obtain (see also [12], p. 239):

$$E(w_{\tau,3}) \simeq 0.628\tau, \quad M_2(w_{\tau,3}) \simeq \frac{2.751}{\tau}, \quad m_2(w_{\tau,3}) \simeq \frac{8.485}{\tau^2}, \quad r(w_{\tau,3}) \simeq \frac{4.378}{\tau^2}, \quad (55)$$

$$E(p_{\tau}) \simeq 0.539\tau, \quad M_2(p_{\tau}) = \frac{3}{\tau}, \quad m_2(p_{\tau}) = \frac{12}{\tau^2}, \quad r(p_{\tau}) \simeq \frac{5.566}{\tau^2}, \quad (56)$$

$$E(a_{\tau}) \simeq 0.587\tau, \quad M_2(a_{\tau}) \simeq \frac{2.79}{\tau}, \quad m_2(a_{\tau}) \simeq \frac{9.87}{\tau^2}, \quad r(a_{\tau}) \simeq \frac{4.753}{\tau^2}. \quad (57)$$

As can be seen from (55) – (57), our window is better than Parzen and Papoulis window with respect to all stated design criteria. However, it should be mentioned that the system functions of Parzen and Papoulis window are nonnegative while this is not the case for our window. This must be taken into account in view of the amplitude moment results. Moreover, it should be noted that $W_{\tau,3}$ has its first null already at $\omega \simeq \frac{2.4226\pi}{\tau}$ while Parzen resp. Papoulis window have their first zero of the system function at $\omega = \frac{4\pi}{\tau}$ resp. at $\omega = \frac{3\pi}{\tau}$. Together with the fast asymptotic decay of $W_{\tau,3}$ this small mainlobe-width should be of some interest in practice (see also Figure 4, and the plots in [10] for further qualitative comparison).

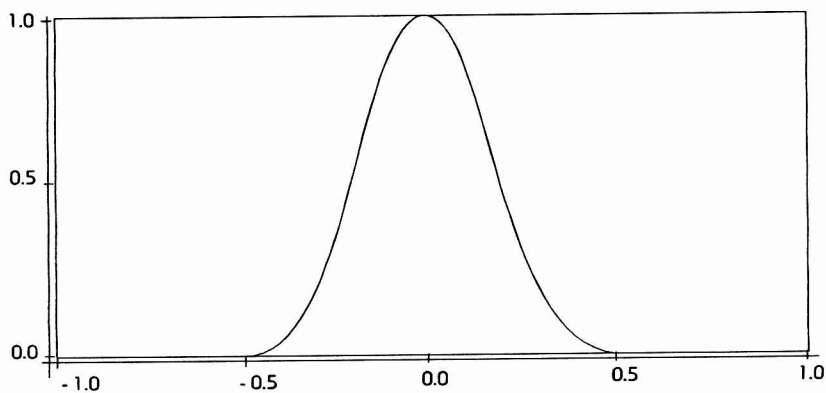


Figure 3. Plot of $w_{r,3}(t)$ for $r = \frac{1}{2}$ and $t \in [-1, 1]$.

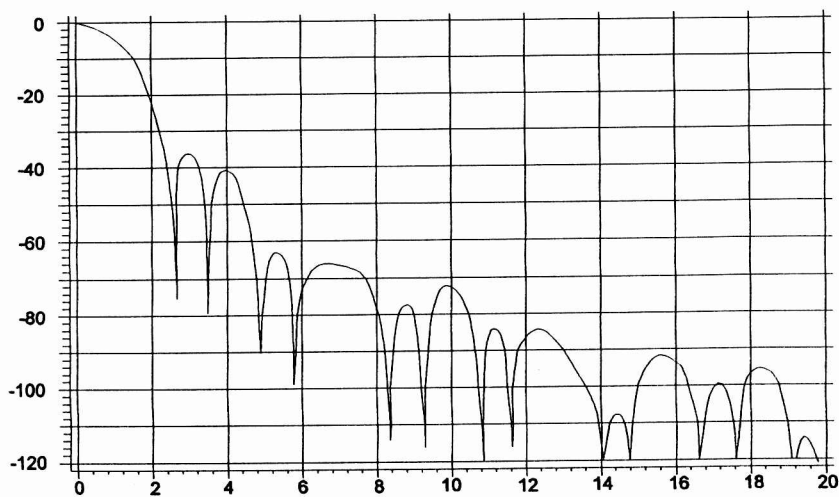


Figure 4. Plot of $W_{r,3}^{dB}(2\pi\omega)$ for $r = \frac{1}{2}$ and $\omega \in [0, 20]$.

5 Conclusions

In this paper, we introduced a family of window pairs which were shown to be optimal in view of a minimization criterion on the highest available derivative of the time-function. This optimal design strategy was motivated by the fact that the derivatives of the window time-function are intimately connected with energy and amplitude moment properties of the frequency-function and controlling them is a standard approach in (optimal) window design. Moreover, we gave an explicit analytical representation of our windows which sometimes is essential in order to avoid sophisticated time-consuming optimal design algorithms. In the following, we took a more detailed look at the first three most relevant windows of our family. In these cases, we explained how the windows could be realized and compared them with well-known extensively used windows as Bartlett, Tukey, Blackman, Parzen, and Papoulis window. Finally, we tried to make sure that our windows are at least as good as the above mentioned standard windows, moreover, with respect to some design criteria they are even better.

References

- [1] Adams, J.W. : *A new optimal window*. IEEE Trans. on Signal Processing **39**, 1753–1769 (1991)
- [2] Deboor, C. : *A remark concerning perfect splines*. Bull. Amer. Math. Soc. **80**, 724–727 (1974)
- [3] Butzer, P.L. and Nessel, R.J. : *Fourier Analysis and Approximation I, One-Dimensional Theory*. Basel 1971
- [4] Geçkinli, N.C. and Yavuz, D. : *Some novel windows and a concise tutorial comparison of window families*. IEEE Trans. Acoust., Speech, Signal Processing **26**, 501–507 (1978)
- [5] Glaeser, G. : *Le Prolongateur de Whitney, Université de Rennes*. Vol. 1, 1966 and Vol. 2, 1967
- [6] Glaeser, G. : *Prolongement extrémal de fonctions différentiables d'une variable*. J. Approx. Theory **8**, 249–261 (1973)
- [7] Karlin, S. : *Some variational problems on certain Sobolev spaces and perfect splines*. Bull. Amer. Math. Soc. **79**, 124–128 (1973)

- [8] **Karlin, S.** : *Interpolation properties of generalized perfect splines and the solutions of certain extremal problems I.* Trans. Amer. Math. Soc. **206**, 25–66 (1975)
- [9] **Lenze, B.** : *On the explicit solution of a time-optimal control problem by means of one-sided spline approximation.* J. Approx. Theory **56**, 297–305 (1989)
- [10] **Nuttall, A. H.** : *Some windows with very good sidelobe behavior.* IEEE Trans. Acoust., Speech, Signal Processing **29**, 84–91 (1981)
- [11] **Oppenheim, A.V.** and **Scharfer, R.W.** : *Discrete-Time Signal Processing.* Englewood Cliffs, New Jersey 1989
- [12] **Papoulis, A.** : *Signal Analysis.* New York 1977
- [13] **Schoenberg, I.J.** : *The perfect B-splines and a time-optimal control problem.* Israel J. Math. **10**, 261–274 (1971)
- [14] **Schumaker, L.L.** : *Spline Functions: Basic Theory.* New York 1981
- [15] **Soliman, S.S.** and **Srinath, D.** : *Continuous and Discrete Signals and Systems.* Englewood Cliffs, New Jersey 1990

received: September 9, 1991

Author:

Prof. Dr. B. Lenze
Fachhochschule Dortmund
Fachbereich Informatik
Postfach 10 50 18
44047 Dortmund
Germany

WOLFGANG ENGEL

Bildungseinrichtungen für mathematisch begabte Schüler in Rostock

Ein Rückblick

Die Pädagogen in der ehemaligen DDR vertraten einen weiten Begabungsbegriff, der außer den Fähigkeiten alle an einer Leistung auf einem bestimmten Gebiet beteiligten Persönlichkeitsqualitäten einbezieht. Wir meinen hier Begabungen von einem hohen Niveau und in einer Richtung, die sich besonders bei der Arbeit im Fach Mathematik äußert.

Hochschullehrer und Assistenten des früheren Rostocker Mathematischen Instituts begannen im Zusammenhang mit der Einführung der Mathematikolympiaden in der DDR 1963 mit einem Kursus für Schüler, die sich bei den Wettbewerben besonders hervorgetan hatten. Beteiligt waren etwa 10 Schüler aus der Stadt Rostock und ihrer Umgebung.

Nachdem das damalige Staatssekretariat für das Hoch- und Fachschulwesen der DDR im August 1964 Regelungen für die Gründung von Spezialklassen (Stufe 11 und 12) an den Mathematisch-Naturwissenschaftlichen Fakultäten erlassen hatte, beschloß die Rostocker Fakultät im Oktober 1964, eine solche Einrichtung zu gründen. Diese sollte der Förderung mathematisch-naturwissenschaftlich besonders begabter Jugendlicher dienen und sich an den Erfahrungen der Universität Moskau orientieren. Da jedoch die materiellen und personellen Voraussetzungen (z.B. Sicherung der Internatsunterbringung von Minderjährigen) nicht geschaffen werden konnten, wurde der Beschluß nicht realisiert.

Im Jahre 1977 begannen (erneut) Planungen für einen Neubau für die Sektion Mathematik, und damit wurde der Gedanke der Einführung einer Spezialklasse wieder aufgegriffen. Der Bau wurde jedoch nicht begonnen. Aber es gelang, die Abteilung Volksbildung des Rates des Bezirkes für die Schaffung eines Zuges (Klassenstufen 9 bis 12) mathematisch besonders begabter Schüler an der 1. Erweiterten Oberschule der Stadt Rostock zu gewinnen (im folgenden kurz Spezialklassen). Einige Internatsplätze wurden in einem Kinderheim der Stadt

zur Verfügung gestellt. Von den 18 Schülern (darunter 6 Mädchen), die 1981 als erste aufgenommen wurden, waren 15 aus dem Stadtgebiet Rostocks, so daß das angestrebte Ziel der Auswahl von höher begabten Jugendlichen wegen der zu kleinen Population nicht erreicht werden konnte. Die Situation verbesserte sich, als durch ein erweitertes Angebot von Internatsplätzen ein größerer Einzugsbereich möglich wurde.

Von der Sektion Mathematik der Universität wurde ein Mitarbeiter für den Mathematikunterricht delegiert, d.h., er blieb Angehöriger der Universität, war aber voll in den Schulbetrieb eingegliedert. Der Unterricht in Mathematik und Physik erfolgte mit erhöhter Stundenzahl, der Inhalt orientierte sich am Vorgehen in der Heinrich - Hertz - Schule in Berlin. Die „wissenschaftlich - produktive Arbeit“ (Projektarbeit) wurde in mathematisch - naturwissenschaftlichen sowie technischen Sektionen und im Rechenzentrum der Universität durchgeführt. Lehrer der Schule wurden durch Hospitationen vom delegierten Mathematiklehrer auf eine spätere Übernahme des Unterrichts vorbereitet. Weitere Hochschullehrer, Mitarbeiter und Studenten beteiligten sich an der außerunterrichtlichen Beschäftigung der Schüler mit Mathematik und Naturwissenschaften.

In den Jahren 1981 bis 1983 analysierte eine Arbeitsgruppe der Akademie der Pädagogischen Wissenschaften die Erfahrungen der mathematisch - naturwissenschaftlichen Bildungseinrichtungen in der Volksbildung und im Hochschulwesen. Im Anschluß daran wurden einheitliche Lehrpläne und Prüfungsanforderungen für den obligatorischen Unterricht erarbeitet, die ab Schuljahr 1986/87 für die neu konzipierten Spezialschulen Geltung bekamen. Damit erhielten die Spezialklassen (SKL, d.h. Züge an einer Polytechnischen Oberschule oder an einer Erweiterten Oberschule) und Spezialschulen (SPS) mathematischer (-physikalischer) bzw. physikalisch - technischer Richtung nach langen Diskussionen eine einheitliche Konzeption. 1990 bestanden 14 Schulen dieser Art: in Rostock, Berlin, Kleinmachnow, Frankfurt, Cottbus, Magdeburg, Halle, Leipzig, Riesa, Dresden, Chemnitz, Jena, Erfurt und Ilmenau. Sie sollten hochbegabte Schüler in den Klassen 9 bis 12 „in besonders effektiver Weise auf ein Hochschulstudium der Mathematik, der Naturwissenschaften oder der Technik - Wissenschaften vorbereiten und sie befähigen, Spitzenleistungen auf diesen Gebieten zu erreichen“. Eine Verkürzung der Schulzeit war nicht erklärtes Ziel dieser Schulen, wurde aber in Einzelfällen ebenso wie der vorzeitige Übergang in eine höhere Klasse ermöglicht (beim Vergleich mit den Schulen in den alten Bundesländern ist aber zu bedenken, daß in der DDR und in den neuen Bundesländern das Abitur generell nach 12 Jahren abgelegt wurde bzw. wird). Diese Schulen waren fast alle den Räten der Bezirke bzw. dem Magistrat von Berlin direkt unterstellt und wurden personell, finanziell und materiell besser ausgestattet als die „Normalschulen“. Im Jahre 1988 wurden ca. 500 Schüler in die Spezialschulen mathematisch - naturwissenschaftlich - technischer Richtung aufgenommen, das waren ca. 0,25 % der Schüler des Jahrgangs.

Zur Ergänzung sei angeführt, daß es 1983 in der DDR

- 6 SPS und 2 SKI mathematischer und physikalisch – technischer Richtung,
- 4 SKI für Mathematik bzw. Chemie an Universitäten und Hochschulen,
- 1 SPS und 49 SKI mit verstärktem neusprachlichem Unterricht,
- 9 SKI mit verstärktem altsprachlichem Unterricht,
- 1 SPS und 8 SKI für Musikerziehung,
- 2 SPS und 1 SKI zur Vorbereitung eines Studiums als Russischlehrer,
- 1 SKI für Musik und 3 SKI an Hochschulen für Musik,
- 3 SKI an künstlerischen Fachschulen für Bühnentanz und Artistik,
- 15 Kinder- und Jugendsportschulen

gab. Die künstlerischen und sportlichen Spezialeinrichtungen verlangten das Vorhandensein einer entsprechenden speziellen Begabung. Das war bei den sprachlichen Einrichtungen nicht immer notwendig.

Die Anzahl der Schüler in den allgemeinbildenden Schulen betrug 1986 1 987 000, in den Erweiterten Oberschulen (Klassen 11 bis 12) 44 000. Damit kamen über die Erweiterte Oberschule ca. 10 % eines Jahrgangs zum Abitur. Weitere ca. 10 000 Schüler legten in den Einrichtungen der Berufsausbildung (Betriebsberufsschulen) das Abitur ab, so daß etwa 12 – 15 % der Schüler eines Jahrgangs die Hochschulreife erlangten.

An 5 Universitäten und Hochschulen bestanden 1983 noch Spezialklassen (Stufen 11 und 12), nämlich für Mathematik und Physik 1 SKI an der Humboldt – Universität in Berlin, 2 SKI an der Martin – Luther – Universität in Halle, 1 SKI an der Technischen Universität in Magdeburg, 2 SKI an der Technischen Universität in Chemnitz und für die Chemie 4 SKI an der Technischen Hochschule Leuna – Merseburg.

Als Ziel des Unterrichts wurde eine hohe Allgemeinbildung und in den „profilbestimmenden Fächern“ (Mathematik, Informatik, Physik, Chemie und Biologie) eine erweiterte und vertiefte Bildung formuliert. Es kam nicht auf eine Verkürzung der Ausbildungszeit an, sondern auf ein tieferes Eindringen und Verarbeiten des angeeigneten Wissens. Die Schüler bekamen vielfältige Möglichkeiten, in hohem Maß kreative Leistungen bei der Lösung von Problemen zu erbringen. Die Inhalte und die Methoden für den Unterricht waren so gewählt, daß die Schüler in starkem Maße zur selbständigen Arbeit angehalten wurden und daß sie dabei tiefer in die Methoden und Erkenntnisverfahren der jeweiligen Fachwissenschaft eingeführt

wurden. Es wurde angestrebt, daß die Abiturienten eines der profilbestimmenden Fächer studieren und sehr zeitig in die jeweilige Forschung einbezogen werden. In einigen Fällen konnten die Abiturienten sofort die ersten Zwischenprüfungen ablegen und somit das erste Studienjahr überspringen.

Auffällig war, daß es nicht gelungen ist, eine angemessene Zahl Mädchen in die Spezialschulen aufzunehmen. So waren 1988/89 in 23 erfaßten zwölften Klassen mit 363 Schülern nur 72 Mädchen (20 %).

Aus dem Mathematik – Zug der Rostocker Erweiterten Oberschule wurde 1986 die Spezialschule mathematisch – naturwissenschaftlich – technischer Richtung des Bezirkes Rostock, die 1990 den Namen „Albert – Einstein – Schule“ erhielt. Ihr Einzugsgebiet waren die Bezirke Rostock, Schwerin und Neubrandenburg, also das heutige Mecklenburg – Vorpommern. Die Rostocker Spezialschule war 1990 mit den Klassenstufen 9 bis 12 vollständig dreizügig ausgebaut. Zur Unterstützung ihrer Arbeit für die Erreichung des angestrebten Zieles schlossen der Rat des Bezirkes, die Universität Rostock, die Ingenieurhochschule Warnemünde – Wustrow und das Kombinat Schiffbau Rostock im Mai 1986 einen Vertrag, der die Zusammenarbeit regeln sollte. Schwerpunkte waren für die Universität die Unterstützung beim Unterricht in Mathematik, Informatik, Physik, Chemie, Biologie und Technik sowie bei der wissenschaftlich – praktischen Arbeit (Projektarbeit). Das Interesse der Universität war durch die Sorge um die Gewinnung eines hochbefähigten Nachwuchses begründet, wobei vor allem die Erfahrungen von Rostocker Mathematikern eine Rolle spielten.

Im November 1989 hatte die Schule 182 Schüler in 11 Klassen (141 Schüler wohnten in zwei Internaten; 46 aus dem Bezirk Rostock, 46 aus dem Bezirk Neubrandenburg, 49 aus dem Bezirk Schwerin). Sie wurden von 46 Pädagogen unterrichtet und in den Internaten betreut. Dazu gab es noch 46 technische Mitarbeiter. Der Etat 1990 umfaßte 1 140 000 M (alle Geldbeträge in Mark der DDR), darunter 713 000 M für Löhne und Gehälter. Die materielle Ausstattung wurde im Zuge des Aufbaus besonders vorangetrieben. Außer der Planung der Lehrmittel in Höhe von 1 000 M / Klasse (300 M / Klasse für die Erweiterten Oberschulen) konnte die Rostocker Spezialschule noch Sondermittel in Anspruch nehmen, z.B. für die Anschaffung von vorrangig gelieferten Computern der DDR – Produktion. 1990 waren für Lehrmittel 14 000 M vorgesehen und zusätzlich 40 000 M Sondermittel beantragt.

Die Pflichtstundenzahl eines Lehrers an der Polytechnischen Oberschule betrug 23 Stunden. Wenn er mehr als 50 % in den Klassen 9 und 10 unterrichtete, nur 22 Stunden. Eine weitere Abminderungsstunde gab es für Klassenleiter. Stellvertretende Direktoren und Fachberater unterrichteten nur ca. 12 Stunden. Das Anfangsgehalt eines Lehrers war 1 000 M, das nach je 2 Jahren bis zu 1 500 M erhöht wurde. Dazu gab es Zuschläge für die Titel Oberlehrer, Studienrat, Oberstudienrat und in Einzelfällen Professor sowie Prämien. Ein Lehrer der Erweiterten Oberschulen erhielt einen Zuschlag von 50 M und hatte 22 Pflichtstunden.

An der Rostocker Spezialschule wurden alle Lehrer wie Lehrer der Erweiterten Oberschule bezahlt. Die Schule hatte überdurchschnittlich viele Möglichkeiten zur Titelvergabe. Die Lehrer der profilbestimmenden Fächer erhielten zur Vorbereitung ihres Unterrichts 4 bis 6 Abminderungsstunden. Einer 2stündigen Arbeitsgemeinschaft waren 3 Stunden Unterricht äquivalent.

In den Rostocker Spezialklassen bzw. in der Spezialschule standen folgende Unterrichtsstunden (in Gesamtwochenstunden für die nachstehend genannten Klassen) zur Verfügung:

	Kl. 9 – 10			Kl. 11 – 12		
	Normal	SKI	SPS	Normal	SKI	SPS
Mathematik	9	16+2	11,5	10	14	11
Informatik	0	0	4	0	0	0
Physik	6	6	7,5	6	8	11
Chemie	4	4	4	5	5	6
Biologie	4	4	4	5	5	3
Einführung in die soz. Produktion	4	3	4	0	0	0
Produktive Arbeit	6	0	6	0	0	0
Wahlweise – obligatorischer Unterricht						
– Math., Inf., Phys., Chem., Bio.	0	(2)	4	(6)	(2)	10
– wissenschaftlich – praktische Arbeit	0	0	0	4	4	4
Russisch	6	6	5	8	7	7,5
Englisch	(5)	5	4,5	(6)	5,5	5,5
Deutsch	7	7	5,5	7	7	6
Staatsbürgerkunde	3	3	2,5	3	3	3
Geschichte	4	4	3	3	3	3
Geographie	3	3	2,5	2	2	0
Sport	4	4	4	4	4	4
Musik/Kunsterziehung	2/1	2/1	2	(1)	2	0
Astronomie	1	1	0	0	0	0
Wochenstundenza. 1en	32	35–36	37	35–36	34–35	37
	+ (2,5)	+ (1)			+ (1)	

Die Zahlen in Klammern bedeuten fakultativen Unterricht.

Der Unterricht in den Sprachen, in Mathematik und den Naturwissenschaften wurde zum Teil in Gruppen von ca. 8 Schülern erteilt. In der Spezialklasse bedeuten die +2 Stunden Mathematik „Numerische Mathematik“ (ein Vorläufer der Informatik), die 2 Stunden fakultativer Unterricht waren Stochastik bzw. Elektronik / Chemie.

Durch kleine Schüleranzahlen (14 – 18) konnte ein sehr intensiver Unterricht erteilt werden, so daß die in den Sprachen etwas reduzierte Stundenzahl beim Vergleich mit dem „Normalschüler“ nicht ins Gewicht fiel.

Die äußere Differenzierung war vor allem auf die Klassen 11 und 12 konzentriert, wo der wahlweise – obligatorische Unterricht in den sog. profilbestimmenden Fächern vornehmlich durch Hochschullehrer und Mitarbeiter der Universität erteilt wurde. Mit einigen hochbegabten Schülern der Klassen 11 und 12 erfolgte eine Einzelförderung durch Hochschullehrer, insbesondere aus den Sektionen Mathematik, Chemie, Informatik und Technische Elektronik. In den Klassen 9 und 10 wurde der wahlobligatorische Unterricht durch Lehrer der Schule bestritten. Einige dieser Lehrer hatten zuvor im Rahmen einer Aspirantur an der Universität promoviert, und es gab einen Plan, nach dem die aus Altersgründen ausscheidenden Lehrer durch postgradual qualifizierte und teilweise promovierte ersetzt werden sollten.

Schüler, die sich besonders für Mathematik interessierten, wurden in die Rostocker Mathematische Schülergesellschaft (MSG) aufgenommen. Diese war eine von der Sektion Mathematik der Universität Rostock und der Abteilung Volksbildung des Rates des Bezirkes gemeinsam getragene Einrichtung. Die Universität sicherte die Veranstaltungen inhaltlich, die Volksbildung materiell und finanziell. Es wurden durch die MSG über die o.a. Studentafel hinaus noch ca. 2 Wochenstunden und jährlich 4 bis 5 Wochenendlehrgänge (Freitagmittag bis Sonntagmittag, häufig in Jugendherbergen o.ä. außerhalb Rostocks, vor den Wettbewerben der Mathematikolympiade) durchgeführt. In die Mathematische Schülergesellschaft wurden auch Schüler anderer Schulen des westlichen Teils des Bezirkes Rostock aufgenommen.

Der Inhalt solcher Veranstaltungen wird am Beispiel des Schuljahres 1984/85 beschrieben (ungefähre Stundenzahl in Klammern).

Im ersten Halbjahr wurden in Vorbereitung der Kreis- und Bezirksolympiade vorwiegend Aufgaben behandelt, nämlich:

In Klasse 9: Einführung zu Zahlenkongruenzen (6), vollständige Induktion (4), Vektorrechnung (6), Ungleichungen, arithmetisches und geometrisches Mittel (4). Neue Schwerpunkte beim Aufgabentraining waren Kombinatorik (Schubfachprinzip, Anzahlen), Extremwertaufgaben ohne Differentialrechnung, Gleichungssysteme (lineare, nichtlineare, symmetrische), Ungleichungen, geometrische Konstruktionen.

In Klasse 10: Einführung zur Koordinatengeometrie (6), Diophantische Gleichungen (6), Funktionalgleichungen (6). Neue Schwerpunkte beim Aufgabentraining waren Trigonometrie und Goniometrie, Ungleichungen mit Beträgen, Logarithmen, Beweise in der Geometrie.

In Klassen 11: Einführungen zu den Gebieten komplexe Zahlen mit geometrischen Anwendungen (4), Polynome (4), Zahlenfolgen (6), abzählende Kombinatorik (6). Neue Schwerpunkte beim Aufgabentraining waren Polynome, Wahrscheinlichkeitsrechnung, räumliche

Geometrie.

In Klasse 12 wurde die Jensensche Ungleichung eingeführt (6). Im Aufgabentraining wurden Färbungen auf Graphen behandelt.

Das zweite Halbjahr diente dem Vortrag mathematischer Themen, die den Schulstoff wesentlich vertiefen und auch erweitern. Im Schuljahr 1984/85 wurden behandelt:

In Klasse 9: Graphentheorie (14), Statistik I (12), Gruppentheorie (12).

In Klasse 10: Kombinatorik und Informatik (10), Numerik und Rechentechnik I (12), Differentialgleichungen (8), Blockpläne und lateinische Quadrate (6).

In Klasse 11: Lineare Optimierung (19), Statistik II (12), Kombinatorische Algorithmen (8), Zahlentheorie (8).

In Klasse 12: Mathematik im Schiffbau (10), Numerik und Rechentechnik II (19), Einführung in die Funktionalanalysis (10), Geschichte der Mathematik (6).

Diese Veranstaltungen sollten den Schülern zeigen, womit sich Mathematiker –i.bes. an der Rostocker Universität– heute befassen. Während des o.a. Schuljahres wurden noch einige Vorträge gehalten, die nicht nur mathematische Gegenstände i.e.S. behandelten, z.B. zur Geschichte der Universität, zu Fragen der Mikroelektronik. An den Veranstaltungen beteiligten sich 5 Professoren, 3 Dozenten, 9 wissenschaftliche Mitarbeiter oder Aspiranten und 4 Studenten. Die jüngeren waren i.d.R. selbst erfolgreiche Olympiade – Teilnehmer und Mitglieder des Bezirksklubs Junger Mathematiker (ein Vorläufer der MSG) gewesen. Um auch schon in den Klassen 5 und 6 die außerunterrichtliche Beschäftigung mit der Mathematik stärker anzuregen, wird seit 1985 von der Universität Rostock für die Schüler Rostocker Schulen im Frühjahr eines jeden Jahres ein Klausurwettbewerb in zwei Runden organisiert, bei dem die erfolgreichste Schule den „Pokal des Rektors“ erhält. Die besten Schüler bekommen Buchprämien.

Die Auswahl der Schüler für die Rostocker Spezialschule (und auch für die anderen SPS) erfolgte über eine Prozeßdiagnostik, wobei neben den Urteilen der Lehrer die der Leiter der außerunterrichtlichen Arbeitsgemeinschaften usw. (also außer besonders befähigten Lehrern auch einschlägige Fachleute von Hochschulen) eine Rolle spielten. Für die Spezialschule wurden die Schüler auf Grund besonderer Leistungen bei Mathematikolympiaden und in verschiedenen außerunterrichtlichen Arbeitsgemeinschaften von den Heimatschulen vorgeschlagen. Aber die Eltern konnten sich auch über die Kreisschulräte direkt an die Spezialschule wenden und ihren Wunsch zum Schulbesuch begründen. Für den Einzugsbereich der Rostocker Spezialschule ist nicht bekannt geworden, daß fachlich leistungsfähige Schüler durch die Kreisschulräte zurückgestellt wurden. Es gab ca. dreimal soviel Bewerber, als Plätze vorhanden waren. Dabei ist natürlich zu bedenken, daß ab 1983 die Dauer der Hochschulvorbereitung auf die Klassen 11 und 12 beschränkt wurde, während die Spezialschule

mit Klasse 9 begann. Für die endgültige Auswahl haben sich dabei das frühzeitige Kennenlernen (etwa bei Mathematikolympiaden u.ä.) und eine längerfristige Zusammenarbeit als zweckmäßig erwiesen. Dazu wurden dreitägige sog. Sichtungslager für Schüler der Klassen 7 und 8 und ggf. Klausuren in den Schwerpunktfächern sowie ein Eignungsgespräch an der Spezialschule durchgeführt. Zwischen den Sichtungslagern gab es noch drei dreitägige Spezialistenlager für Mathematik, Physik und Chemie. Es bestand auch die Möglichkeit, über Korrespondenzaufgaben brieflich mit der Schule in Verbindung zu bleiben. Weiterhin führten Mitarbeiter der Sektion Pädagogik und Psychologie Tests (IST) durch, die aber für die Aufnahme in die Schule keine wesentliche Rolle spielten.

Bei der Aufnahme der Schüler in die Rostocker Spezialschule wurde erwartet, daß die Schüler die Normen der Staatsbürger der DDR einhalten würden. Das wurde aber bei begabten Schülern sehr großzügig gefaßt. Konfessionelle Bindungen wurden weder erfragt noch irgendwie registriert. Eine gewisse Voraussetzung für die Aufnahme in die Spezialschule war aber der Schülerwunsch, ein Studium in Richtung Mathematik, Informatik, Naturwissenschaften oder Technik aufnehmen zu wollen (in Einzelfällen wurden nach dem Abitur auch andere Studien begonnen).

Die Ergebnisse im Abitur waren gut bis sehr gut und trotz der erhöhten Anforderungen besser als in den anderen „normalen“ Schulen. Preise bei DDR – Wettbewerben und ein 2. Preis bei der Internationalen Chemieolympiade wurden im Bericht der Schulleitung 1989 hervorgehoben. Allerdings ist zu bedenken, daß nur vier Abiturientenjahrgänge der Rostocker Spezialklasse und zwei Jahrgänge der Spezialschule bewertet werden können.

Die Spezialschulen der DDR hielten enge Kontakte und tauschten ihre Erfahrungen aus. Es wurden z.B. gesonderte Weiterbildungslehrgänge für die Spezialschullehrer durchgeführt. Auch mit ähnlichen Einrichtungen in den ehemaligen sozialistischen Ländern wurden Verbindungen geknüpft. Für die Rostocker Schule gab es ab 1988 sogar Kontakte zu der Jugenddorf – Christophorusschule in Braunschweig, der damals einzigen Schule in der alten Bundesrepublik, die Hochbegabtenklassen führte. Die dort gemachten Erfahrungen wurden im Rostocker Lehrerkollegium diskutiert. Eine Folge war die Einstellung einer Psychologin. Da die meisten Materialien über die Spezialschulen vertraulich („nur für den Dienstgebrauch“ oder mehr) waren und ein Besuch der Spezialschule durch Nicht – DDR – Bürger die Genehmigung der Ministerien erforderten, war die wissenschaftliche Diskussion der Probleme sehr schwierig.

Bei der Bewertung der Spezialschulen muß man von ihrer Einbettung in das Schulsystem und damit in das Gesellschaftssystem des jeweiligen Staates ausgehen. Diese waren in der DDR durch eine ideologische Überfrachtung, durch einen übertriebenen Zentralismus und Dirigismus, durch Unbeweglichkeit und Angst vor Machtverlust gekennzeichnet. Aber allgemein gilt, daß ein Schulwesen nicht allein durch Verordnungen und Lehrpläne bestimmt wird, sondern durch die Lehrer und Schulleiter, die die Ordnungen auszulegen haben. Deshalb

sollte man würdigen, was viele trotz der komplizierten Bedingungen geleistet haben. Da die schriftlichen Abituraufgaben in Mathematik für alle Spezialschulen von einer zentralen Arbeitsgruppe vorbereitet und (ohne Information des Ministeriums) ausgewertet wurden, war an den Ergebnissen deutlich zu erkennen, welche Schulen ihre Schüler (und vielleicht auch Lehrer) nicht gründlich ausgewählt hatten. Doch waren unter den Bedingungen der DDR die Spezialschulen eine besonders effektive und erfolgreiche Form der Begabtenförderung und Hochschulvorbereitung. Die durchgeführten Untersuchungen ergaben, daß i.a. keine zu enge Spezialisierung vorhanden war, sondern ein breites Interessenspektrum. Die allgemeinen Bedingungen der DDR und des Ortes ließen aber nicht die Befriedigung aller Interessen in dem vielleicht gewünschten Maße zu. Da auf Grund der Auswahl der Schüler die Spezialschulen auch Schüler hatten, deren Eltern nicht am Schulort wohnten, mußten sie bis auf Ausnahmen (z.B. Berlin, Dresden) Internate besitzen. Das muß kein Nachteil sein, weil historische Erfahrungen belegen, daß Internatsschulen hohe erzieherische Potenzen besitzen können (z.B. Schulpforte, Franckesche Stiftungen).

Nach der Wiedereinführung der Länder wurde die Sonderstellung der Spezialschulen aufgehoben. Sie wurden i. wes. Gymnasien mit mathematisch - naturwissenschaftlichem Profil. Die Rostocker Schule wurde zusammen mit der Kinder- und Jugendsportschule vom Christlichen Jugenddorfwerk Deutschlands als Christophorusschule übernommen (Ganztagsschule, Klassen 5 bis 12, entsprechend den in Mecklenburg - Vorpommern geltenden staatlichen Regelungen). Die Klassenstärke beträgt ca. 22 Schüler. Die Sprachenfolge ist Englisch (ab Kl. 5), Russisch, Latein, Französisch (ab Kl. 7); Latein, Französisch, Dänisch oder Schwedisch (ab Kl. 9). Der Unterricht hat einen Schwerpunkt im mathematisch - naturwissenschaftlichen Bereich. In Mathematik, Physik und Chemie gibt es im Rahmen der staatlichen Anerkennung etwas erhöhte Stundenzahlen. Weiter wird ein Informatik - Technik - Unterricht in den Klassen 5 bis 11 durchgeführt. Leistungskurse werden in allen Naturwissenschaften angeboten. In einem Zug werden nach Braunschweiger Vorbild ab Klasse 9 intellektuell Spitzenbegabte besonders auf mathematisch - naturwissenschaftlichem Gebiet begleitet. Ebenfalls nach Braunschweiger Vorbild sind Arbeitsgemeinschaften mannigfacher Art vorgesehen, darunter natürlich auch solche für Mathematik und Naturwissenschaften. In Zusammenarbeit mit der Universität gibt es weiterhin eine Spitzen- bzw. Einzelförderung in bestimmten Fächern. Ein weiterer Bildungsgang wird für leistungssportlich orientierte Schüler in Zusammenarbeit mit Rostocker Sportvereinen angeboten (Schwerpunkte Schwimmen, Turnen, Wasserspringen und Leichtathletik).

Die bisherige erfolgreiche Arbeit in der Spezialschule und in der Kinder- und Jugendsportschule zeigt sich darin, daß bei mathematisch - naturwissenschaftlichen und sportlichen Wettbewerben die Rostocker Christophorusschule in Mecklenburg - Vorpommern -und wie die Mathematikolympiade 1993 zeigte, darüber hinaus- und im Rahmen des Christlichen

Jugenddorfwerkes eine Spitzenstellung einnimmt.

Damit scheint die in Rostock bestehende Tradition der Förderung mathematisch begabter Jugendlicher auch für die Zukunft gesichert zu sein.

Literatur

- [1] Engel, Wolfgang : *Entdeckung und Förderung mathematischer Begabungen in der DDR*. Zentralblatt Didaktik d. Mathematik 1990, Heft 1, 23-34
- [2] Pollmer, K. : *Information über einige empirische Ergebnisse zum Begabungsniveau und zur Begabungsrichtung von Spezialschülern mathematisch – naturwissenschaftlich – technischer Richtung*. Leipzig, Zentralinstitut für Jugendforschung 1989. (Dieses Informationsmaterial war für die Leitungen der Spezialschulen bestimmt.)

Unser früherer Mitarbeiter Dr. Wolfgang Guba (jetzt am Gymnasium in Grevesmühlen), der die erste Spezialeklasse vier Jahre als Mathematiklehrer begleitet hat, und Oberstudienrat Klaus Krüger, der frühere Direktor der Spezialschule Rostock, haben diesen Beitrag kritisch durchgesehen und ergänzt, wofür ich ihnen danke.

eingegangen: September 9, 1993

Verfasser:

Prof. Dr. W. Engel
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

Hinweise für Verfasser

Um die redaktionelle Bearbeitung und die Herstellung der Druckvorlage zu erleichtern, wären wir den Verfassern dankbar, sich betreffs der Form der Manuskripte an den in **Rostock. Math. Kolloq.** (ab Heft 43) veröffentlichten Beiträgen zu orientieren. Insbesondere beachte man:

1. Manuskripte sollten grundsätzlich **maschinengeschrieben** (Schreibmaschine, Drucker) in **deutscher oder englischer Sprache** abgefaßt sein.
2. Zur inhaltlichen Einordnung der Arbeit sind **1–2 Klassifizierungsnummern** (entsprechend der "1980 Mathematics Subject Classification" der Mathematical Reviews) anzugeben.
3. **Textbreite/Texthöhe** des Manuskripts sollten sich an den Maßen **160mm/230mm** orientieren.
4. Der Manuskripttext ist **eineinhalbzeilig, linksbündig**, wenn möglich, links- und rechtsbündig zu schreiben. Beim Auftreten von Formeln im laufenden Text ist der Zeilenabstand entsprechend zu vergrößern.
5. Der Platz für **Abbildungen** sollte beim Schreiben **ausgespart** werden; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert beizufügen.
6. **Literaturzitate** sind im Text durch laufende Nummern (vgl. [3], [4] ; [7, 8, 10]) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift **Literatur** bzw. **References** zusammenzustellen. Hierbei ist die durch die nachfolgenden Beispiele veranschaulichte Form einzuhalten (die Zeitschriftenabkürzungen erfolgen nach Mathematical Reviews).

[3] Zariski, O., and Samuel, P.: *Commutative Algebra*. Princeton 1958

[4] Steinitz, E.: *Algebraische Theorie der Körper*. J. Reine Angew. Math. **137**, 167–309 (1920)

[8] Gnedenko, B.W.: *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung*. In: Reichard, H. (Ed.): C.F. Gauß, Gedenkband anlässlich des 100. Todestages. S. 193–204, Leipzig 1967

Die Angaben erfolgen in Originalsprache ; bei kyrillischen Buchstaben sollte die bibliobekarische Transkription bzw. eine Übersetzung lt. Mathematical Reviews verwendet werden.

7. Die **aktuelle, vollständige Adresse des Verfassers** sollte enthalten: **Titel / Vornamen Name / Institution / Struktureinheit / Straße Hausnummer / Postleitzahl Ort / Land**.

Weiterhin besteht die Möglichkeit, mit dem Satzsystem **TEX** erstellte Manuskripte auf unter **MS-DOS** formatierten Disketten (**3.5"**, **0.72MB**, **1.44MB**; **5.25"**, **0.36MB**, **0.72MB**, **1.2MB**) einzureichen.

