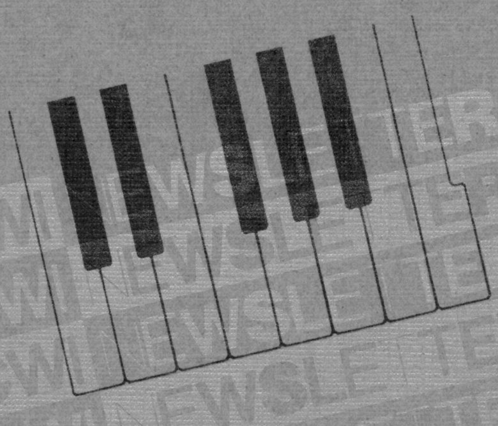
CWI NEWSLETTER

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

Quarterly, Issue no. 13
December 1986



CWI NEWSLETTER

Number 13, December 1986

Editors

Arjeh M. Cohen Richard D. Gill Jo C. Ebergen

Executive editor

Wim A.M. Aspers

Editorial secretary

Wilma E.G. van Eijk

The CWI Newsletter is published quarterly by the Centre for Mathematics and Computer Science (Centrum voor Wiskunde en Informatica), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands. The Newsletter will report on activities being conducted at the Centre and will also contain articles of general interest in the fields of Mathematics and Computer Science, including book reviews and mathematical entertainment. The editors encourage persons outside and in the Centre to contribute to the Newsletter. Normal referee procedures will apply to all articles submitted.

The Newsletter is available free of charge to all interested persons. The Newsletter is available to libraries on an exchange basis.

All correspondence should be addressed to: *The CWI Newsletter, Centre for Mathematics and Computer Science, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands.*

ISSN 0168-826x

Contents

Centre for Mathematics and Computer Science
Centrum voor Wiskunde en Informatica

The Centre for Mathematics and Computer Science (CWI) is the research institute of the Stichting Mathematisch Centrum (SMC), which was founded on 11 February 1946.

The goal of CWI is to carry out fundamental and advanced research in mathematics and computer science, with special attention to those areas in which the research may have important applications.

Research at CWI is organized in eight scientific departments:

- Pure Mathematics;
- Applied Mathematics;
- Mathematical Statistics;
- Operations Research and System Theory;
- Numerical Mathematics;
- Software Technology;
- Algorithms and Architecture;
- Interactive Systems

There are also a number of supporting service divisions, in particular the Computer Systems and Telematics Division, and an extensive Library.

The subdivision of the research into eight departments is less rigid than it appears, for there exists considerable collaboration between the departments. This is a matter of deliberate policy, not only in the selection of research topics, but also in the selection of the permanent scientific staff.

Symplectic Automorphisms of K3-Surfaces

(after S. Mukai and V.V. Nikulin)

Geoffrey Mason
University of California at Santa Cruz

A short, but fairly complete, account is given of the work of MUKAI and NIKULIN on so-called symplectic automorphisms of $K3$-surfaces. (Nikulin calls such automorphisms *algebraic.*)

1. INTRODUCTION

If X is a $K3$-surface then, essentially by definition, X has a nowhere-vanishing holomorphic 2-form ω. The group G of automorphisms of X which preserves ω is called the group of *symplectic automorphisms* of X, and the combined work of MUKAI [7] and NIKULIN [8] gives a complete list of the possibilities for the isomorphism type of G (it is known that G is a finite group). In fact, one can be much more precise, in particular Mukai shows that there is an imbedding $i: G \rightarrow M_{23}$ where M_{23} is one of the sporadic Mathieu groups (see below) and that this induces a $\mathbb{Q}G$-isomorphism from the total rational cohomology $V = H^*(X, \mathbb{Q})$ of X to the usual permutation module P (of degree 24) of M_{23}. Furthermore by Hodge theory one knows that $dim V^G \geqslant 5$ (V^G = the space of G-invariants in V), so that because of the isomorphism $V \cong P$ we find that $i(G)$ has at least five orbits on the 24 letters being permuted. Mukai shows that, conversely, if $H \leqslant M_{23}$ has at least five orbits on the 24 letters then there is a $K3$-surface on which H acts (effectively) as symplectic automorphisms. Because of the surprising connection with M_{23}, these results are of interest to finite group-theorists as well as others.

I have taken the opportunity of simplifying Mukai's group-theoretic analysis of the possibilities for G. Thus only some standard facts, available in [6] and the elementary parts of [4], are needed, and the only *classification* results we use are the results of BRAUER [2] giving the simple groups of order $2^a \cdot 3^b \cdot 5$ (needed only if $a \leqslant 7$, $b \leqslant 2$!). There is almost nothing new here, although we should comment that for the possibility $G \cong \mathbb{Z}_2 \times \mathbb{Z}_8$ Mukai has to refer to Nikulin, whose proof that this cannot occur is quite intricate. In fact we will eliminate this possibility quite easily at the outset, making the overall proof reasonably short.

2. $K3$-SURFACES

We recall some pertinent facts about $K3$-surfaces. $K3$-surfaces were named (by A. Weil) after Kummer, Kähler, Kodaira and the beautiful mountain $K2$ in Kashmir. They form one of 10 classes of minimal models of compact connected 2-dimensional complex manifolds in the Enriques-Kodaira classification (and one of 5 classes of such manifolds with Kodaira dimension 0). Double coverings of the complex projective plane with a branch curve of degree 6 having only simple singularities are examples, but there are $K3$-surfaces that cannot be constructed in this way. (In fact, the set of algebraic $K3$-surfaces is a union of countably many 19-dimensional families in the 20-dimensional family of all $K3$-surfaces.) For background material on $K3$-surfaces we refer the reader to [1]. Those readers not familiar with $K3$-surfaces may take the relevant results below as axioms without impairing their understanding of the later group-theoretic analysis.

We may define a $K3$-surface X to be a compact, 2-dimensional complex manifold such that X has first Betti number 0 and trivial canonical bundle. Then the cohomology space $V = H^{*}(X,\mathbb{Q})$ is even, that is the groups $H^{1}(X,\mathbb{Q})$ and $H^{3}(X,\mathbb{Q})$ are trivial. The Hodge decomposition yields a representation of V as a direct sum

$$V = H^{0}(X)\oplus H^{2,0}(X)\times\oplus H^{1,1}(X)\oplus H^{0,2}(X)\oplus H^{4}(X)\,, \tag{2.1}$$

each of the five summands being non-trivial. In fact we have

$$dimH^{0} = dimH^{2,0} = dimH^{0,2} = dimH^{4} = 1 \text{ and } dimH^{1,1} = 20\,. \tag{2.2}$$

In particular $dimV = 24$. Now the group G of symplectic automorphisms of X is the group of automorphisms of X which acts trivially on $H^{2,0}(X)$. Then by duality G is trivial on $H^{0,2}(X)$, as well as being trivial on $H^{0}(X)$ and $H^{4}(X)$. Mukai establishes that $H^{1,1}(X)$ also has a non-zero G-invariant, whence

$$dimV^{G} \geqslant 5. \tag{2.3}$$

3. THE REPRESENTATION OF G ON $H^{*}(X,\mathbb{Q})$

We have already remarked that G is a finite group, and Mukai's method is to first compute the character of G on the rational G-space V. This proceeds as follows: first fix a point $p \in X$, and let $G_p = \{g \in G \mid g{\cdot}p = p\}$ be the corresponding isotropy group. Then G_p acts on the tangent space T_p at p (essentially via the map $g \mapsto dg$ sending an element $g \in G_p$ to its differential). This action is faithful, so realizing G_p as a subgroup of $GL(T_p) \cong GL_2(\mathbb{C})$. But in fact, because G_p preserves the 2-form ω, G_p preserves a non-degenerate symplectic form on T_p, yielding

There is an imbedding $G_p \to SL_2(\mathbb{C})$. In particular, the abelian subgroups of G_p are cyclic. (3.1)

Of course, all finite subgroups of $SL_2(\mathbb{C})$ are well known. Next, one knows that for each $g \in G$, the set $F(g) = \{p \in X | g \cdot p = p\}$ of fixed-points of g is finite. Using a version of the Atiyah-Singer index theorem Mukai proves the following crucial result.

The cardinality $|F(g)|$ *of the g-fixed-points depends only on the order* $|g|$ *of* g *, and is given by the formula* $|Fg| = \frac{24}{n}\prod_{p|n}(1+\frac{1}{p})$. (3.2)

Here, $n = |g| \geqslant 2$, and p ranges over the prime divisors of n. This determines the character of G on V; because of the fact that $H^*(X,\mathbb{Q})$ is even and $F(g)$ is finite, the Lefschetz fixed-point-formula tells us the following:

If $V = H^*(X,\mathbb{Q})$ *affords the character* χ *of* g *then for all* $1 \neq g \in G$ *we have* $\chi(g) = |F(g)|$. (3.3)

We complete this section with the possibilities for $\chi(g)$ which follow from (2.2), (3.2) and (3.3). We may write $\chi(n)$ instead of $\chi(g)$ if g has order n, and we obtain the following

n	1	2	3	4	5	6	7	8	9	11	12	15	16	23
$\chi(n)$	24	8	6	4	4	2	3	2	2	2	1	1	1	1

(3.4)

All we need, to determine the nature of G, are the results (2.3) and (3.1)-(3.4), and as we said these may be taken as axioms in the following since nothing more concerning the nature of χ will be needed.

4. The Mathieu groups M_{23} and M_{24}

We record some results concerning the Mathieu groups. Let Ω be a set of cardinality 24 on which the symmetric group Σ_{24} acts in the usual fashion. Σ_{24} contains a (maximal) subgroup M_{24} (which acts quintuply transitively!) on Ω. It can be defined as the stabilizer in Σ_{24} of a collection of 759 subsets of Ω of size 8 with the property that no two have more than 4 elements of Ω in common, and is closely related to the so-called extended binary Golay code. See [3] for more details. M_{24} is a simple group, as is the subgroup M_{23} which is by definition the isotropy group (in M_{24}) of a point of Ω. Both M_{23} and M_{24} are among the so-called sporadic simple groups, which accounts for the interest of the results to group-theorists. However, both the simplicity and sporadic nature of these groups is irrelevant as regards the present discussion. Exactly why M_{23} plays a rôle is presently unclear.

Next we list those isomorphism types of subgroups of M_{23} which, up to conjugacy, are maximal subject to having at least five orbits on Ω. The results can be readily checked, for example, from table 3 of [3].

(i) $PSL_2(7)(\cong SL_3(2))$
(ii) $A_6(\cong PSL_2(9))$
(iii) Σ_5
(iv) $E_{16} : A_5$ (no elements of order 6)
(v) $E_9 : Q_8$

(vi) $E_9 : D_8$
(vii) $(A_4 \times A_4) : \mathbb{Z}_2$ (a 3-Sylow being inverted by an involution)
(viii) $E_{16} : D_{12}$ (trivial center)
(ix) $(Q_8 \star Q_8) : \mathbb{Z}_3$ (no elements of order 12)
(x) $E_{16} : \Sigma_4$ (no elements of order 6)
(xi) $GL_2(3)(\cong Q_8 : \Sigma_3)$

Here, we have used fairly standard notation: A_n is the alternating group on n letters, Σ_n the corresponding symmetric group, E_{p^r} the elementary abelian group of order p^r for a prime p, Q_8 the quaternion group of order 8 and D_{2k} the dihedral group of order $2k$. Furthermore, if A and B are groups, then $A : B$ denotes a semidirect product with normal subgroup A, and $A \star B$ a central product of A and B (i.e., a quotient of $A \times B$ by a subgroup of its centre). The information provided specifies a unique group in each of (i)-(xi).

We remark that the requirement of having at least five orbits on Ω does not prevent some of the groups listed exhibiting several different orbit structures (though it is evident that the *number* of orbits depends only on the group). Thus M_{23} contains a conjugacy class of $PSL_2(7)$ with orbit lengths 1,1,1,7,14 and another with orbits 1,1,7,7,8.

Denote by P the permutation module for M_{24} obtained from its action on Ω. We regard P as a $\mathbb{Q}H$-module for each subgroup H of M_{24} by restriction. If P affords the character π then of course for $g \in M_{24}$ one has

$$\pi(g) = \#\ \textit{of letters in}\ \Omega\ \textit{fixed by}\ g.$$

It is readily verified that the following holds:

If $g \in M_{23}$ then the value of $\pi(g)$ depends only on the order of g, and is given by the same formula as in (3.2), viz. $\pi(g) = \dfrac{24}{n}\prod_{p|n}(1+\dfrac{1}{p})$ *where* $n = |g|$. (4.1)

5. Mukai's results

It was Mukai who first noticed the strange coincidence of (4.1) and (3.2) and used it to prove the following results:

Theorem (Mukai). *If G is a group of symplectic automorphisms of the $K3$-surface X then there is an imbedding $i : G \rightarrow M_{23}$ such that $i(G)$ is a subgroup of one of the groups listed in* (i)-(xi) *above.*

Theorem. *If $V = H^{\star}(X, \mathbb{Q})$ and P are as above then the imbedding i induces a $\mathbb{Q}G$-isomorphism of modules $\alpha : V \rightarrow P$.*

Remarks

(i) Granted the existence of i, the second theorem is a consequence of (4.1).
(ii) Granted the existence of *any* imbedding $G \rightarrow M_{23}$, α still exists, whence by (2.3) we can conclude that $i(G)$ has at least five orbits on Ω and hence is in one of the groups (i)-(xi). Thus to prove the theorems, we

only need *some* imbedding $G \to M_{23}$.

(iii) Mukai exhibits $K3$-surfaces admitting each of the groups (i)-(xi) as symplectic automorphisms, but we will not deal with that result here.

The remainder of this paper is concerned with a proof of Mukai's theorem. As we already mentioned, we use only the results (2.3), (3.1)-(3.4), together with the following facts concerning M_{23}.

6. Some properties of M_{23}

We list here some more technical results concerning the 2-structure of M_{23}. First we introduce the group $\hat{A}_8$, the (universal) central extension of A_8 by $\mathbb{Z}_2$, i.e., the non-split extension $1 \to \mathbb{Z}_2 \to \hat{A}_8 \to A_8 \to 1$. This group will play a rôle in what follows. For now we need only the following:

$\hat{A}_8$ and M_{23} have isomorphic Sylow $2-$subgroups, of order 2^7. (6.1)

(See [5].) Let T be a 2-Sylow subgroup of M_{23}. As M_{23} has a subgroup of the shape $E_{16} : A_7$, the following properties of T are easily verified.

Let $J = J(T)$ be the subgroup of T generated by all abelian subgroups of maximal order (which order is 16). Then J has order 2^6 and it contains exactly two subgroups of type E_{16}, three of type $\mathbb{Z}_4 \times \mathbb{Z}_4$, and no others of order 16 which are abelian. (6.2)

An immediate consequence is

T has no subgroup isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_8$. (6.3)

7. Nikulin's result

Given our group G of symplectic automorphisms, let $a(n) = \#\{g \in G \,|\, |g| = n\}$. Then of course we have

$$\sum_{n \geqslant 1} a(n) = |G|$$

and it is also well known that, in the notation of (3.3),

$$|G|^{-1} \Sigma a(n) \chi(n) = dim V^G$$

this latter quantity being $\geqslant 5$ by (2.3). For example if $|G|$ is a prime p we get $a(1) = 1$, $a(p) = p-1$, $\chi(1) = 24$, $\chi(p) = 24/p+1$ and find that

$$|G|^{-1} \Sigma a(n) \chi(n) = 48/p+1 \geqslant 5.$$

This forces $p \leqslant 7$, and a similar analysis shows that if G is cyclic of order $n = |G|$ then the inequality $n \geqslant 9$ leads to $|G|^{-1} \Sigma a(n) \chi(n) < 5$. So we get

If $g \in G$ then $|g| \leqslant 8$. (7.1)

Next we prove

If $t \in G$ is an involution then $C_G(t)$ is isomorphic to a subgroup of $\hat{A}_8$. (7.2)

As a corollary, we obtain

A 2-Sylow subgroup of G is isomorphic to a subgroup of M_{23}. (7.3)

The corollary follows from (7.2) and (6.1). To prove (7.2), let $F = F(t)$ be the points of the $K3$-surface X fixed by t, so that $|F| = 8$ by (3.4). Of course the group $C = C_G(t)$ preserves F. Now if an element $g \in C$ fixes each point of F then $|g| = 1$ or 2 by (3.4), whereas by (3.1) no two distinct involutions of G can fix a common point of X. This shows that only $<t>$ fixes each point of F, yielding an imbedding $C/<t> \hookrightarrow \Sigma_8$. Let us write $\bar{C} = C/<t>$, thinking of $\bar{C}$ as a group permuting F.

Suppose that $\bar{C}$ contains an odd permutation. Then it contains a permutation $\bar{x}$ of the shape (12), (1234), (12)(34)(56), (123456), or (12345678). In the first four cases $\bar{x}$ fixes points of F, as does the group $<x,t>$. By (3.1) we get $<x,t>$ cyclic, so $|x|=2|\bar{x}|=4$, 8, 4 or 12, respectively. So the fourth case is out by (7.1). In the first three cases, $\chi(x) = 4$, 2, or 4 by (3.4), whereas since $F(x) \subseteq F$ we see that $|F(x)| = 6$, 4, or 2 respectively. This contradiction shows that $\bar{x}$ can only be an 8-cycle.

Now as G has no elements of order 16 by (7.1) then $<x,t> \cong \mathbb{Z}_2 \times \mathbb{Z}_8$. Let u be the unique involution of $<x,t>$ which has a square root in $<x,t>$, and apply a similar procedure to $C_G(u) = B$, say, setting $E = F(u)$. Of course $<x,t> \leqslant B$. Moreover as $t \neq u$ then t fixes no element of E by (3.1), whereas any element $y \in <x,t>$ of order 8 must satisfy $y^4 = u$, so y fixes just 2 points of E by (3.4), so y has shape (1234)(56) in its action on E as y induces an element of order 4. Thus $<x,t>$ induces a group of even permutations of E, putting $<x,t>/<u> \hookrightarrow A_8$.

Finally, this means either $<x,t> \hookrightarrow \mathbb{Z}_2 \times A_8$ or $<x,t> \hookrightarrow \hat{A}_8$, and since neither $\mathbb{Z}_2 \times A_8$ nor $\hat{A}_8$ contains $\mathbb{Z}_2 \times \mathbb{Z}_8$ (cf. (6.1)-(6.3)), this shows that $\bar{x}$ does not exist.

Thus we get $\bar{C} \hookrightarrow A_8$, so that $C \hookrightarrow \mathbb{Z}_2 \times A_8$ or $C \hookrightarrow \hat{A}_8$. Assume the first case. Then no involution of $\bar{C}$ fixes a point of F by (3.1), so a Sylow 2-subgroup of $\bar{C}$ has order at most 8. Also, $\bar{C}$ cannot contain an element of the shape (123) by (3.4), and it has no elements of order 5 or 7 since there are no elements of order $2p$ for $p \geqslant 5$ by (7.1). Thus $|\bar{C}|$ divides 24 and we easily verify that, whatever the possibility for C, the group $\hat{A}_8$ has a subgroup isomorphic to C. So in any case $C \hookrightarrow \hat{A}_8$ as required.

Let $A \leqslant G$ be an abelian subgroup of even order. Then A is isomorphic to a subgroup of one of the following groups: E_{16}, $\mathbb{Z}_4 \times \mathbb{Z}_4$, $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_6$. (7.4)

PROOF. A contains an involution t, so by (7.2) we get $A \hookrightarrow \hat{A}_8$. If A is a 2-group then we may take $A \leqslant T$, a 2-Sylow of $\hat{A}_8$, in which case one of the first three possibilities applies by (6.2). If A is not a 2-group, it must have order $2^a \cdot 3$ for some a since there are no elements of order $2p$ for primes

$p \geqslant 5$. Similarly there are no elements of order 12, so the 2-Sylow subgroup of A is E_{2^a} and $A \cong E_{2^{a-1}} \times \mathbb{Z}_6$. Finally, setting $\bar{A} = A/\langle t\rangle \leqslant A_8$ as in the proof of (7.2), it was shown there that an element of $\bar{A}$ of order 3 necessarily has the shape (123)(456). Then if $4 \mid |\bar{A}|$ then $\bar{A}$ contains an involution $\bar{b}$ fixing both 7 and 8. Then $\bar{b}$ pulls back to an element of order 4 in $A \leqslant \hat{A}_8$, and as A contains no such element this is impossible. So $4 \nmid |\bar{A}|$, that is $a \leqslant 2$ and (7.4) follows.

We can use similar arguments to that of the proof of (7.2), applied to elements of odd prime order p. If x is such an element then $C_G(x)$ acts on the set $F(x) = F$ of fixed points of x in its action on the surface X. Again the group $C_G(x)/\langle x\rangle$ induces a group of permutations on X. If $p \geqslant 5$ then $|X| < p$ by (3.4), so $p \nmid |C_G(x){:}\langle x\rangle|$, which forces $\langle x\rangle$ to be already a p-Sylow subgroup of G. If $p = 3$ then $|X| = 6$ and $C_G(x)/\langle x\rangle \leqslant \Sigma_6$. In this case no element of $C_G(x)/\langle x\rangle$ of order 3 can be a permutation of the shape (123) (by (3.1) and (7.1)), so $|C_G(x)/\langle x\rangle|$ cannot be divisible by 9, that is $|C_G(x)|$ is not divisible by 3^3. By taking x to be an element in the center of a 3-Sylow subgroup, say R, of G, we deduce that $|R| \leqslant 9$. Hence, we have proved

$|G|$ divides $2^7 \cdot 3^2 \cdot 5 \cdot 7$. (7.5)

(NIKULIN). *Any abelian subgroup of G is contained in one of the following groups:* E_{16}, $\mathbb{Z}_4 \times \mathbb{Z}_4$, $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_6$, E_9, $\mathbb{Z}_5$ or $\mathbb{Z}_7$. (7.6)

PROOF. We are done by (7.4) if the abelian group has even order. If not, it has order dividing $3^2 \cdot 5 \cdot 7$ by (7.5) and each element has order 3, 5, or 7 by (7.1). The result follows.

8. THE CASE WHERE $|G|$ IS DIVISIBLE BY 7

We show in this section

If $7 \mid |G|$ then G is isomorphic to a subgroup of L_2 (7). (8.1)

We fix a 7-Sylow subgroup S of G, so that $|S| = 7$ by (7.5). By (7.1) we see that S is its own centralizer in G, that is $S = C_G(S)$. Set $N = N_G(S)$.

The order of N is odd. (8.2)

PROOF. Otherwise there is an involution $t \in N$, and since t does not centralize S then $\langle S,t\rangle \cong D_{14}$. Setting $H = \langle S,t\rangle$, H contains the identity, 6 elements of order 7 and 7 of order 2. This yields

$$dim V^H = \frac{1}{14}(24 + 6 \cdot 3 + 7 \cdot 8) = 7.$$

On the other hand we also have

$$dim V^S = \frac{1}{7}(24+6\cdot 3) = 6,$$

putting us in the situation that $S \leqslant H$ and yet $dim V^H > dim V^S$. This is impossible, and the result follows.

S normalizes no non-trivial subgroup of G of order coprime to 7. (8.3)

PROOF. For let $1 \neq H \leqslant G$ satisfy $H \triangleleft HS$ and $7 \nmid |H|$. By a well known result [4, Theorem 6.2.] we can take H to be a p-group for some prime p. As $S = C_G(S)$ we must have $7 | \ |H|-1$, so by (7.5) we see that $p = 2$. Moreover S acts on $Z(H)$, and by (7.6) together with $7||Z(H)|-1$ we find that $Z(H) \cong E_8$. Set $K = Z(H)\cdot S$. Then K contains 48 elements of order 7 and 7 involutions, yielding

$$dim V^K = \frac{1}{56}(24+7\cdot 8+48\cdot 3) = 4.$$

This contradicts (2.3), and completes the proof.

If G is solvable then $G \cong \mathbb{Z}_7$ *or* $\mathbb{Z}_7{:}\mathbb{Z}_3$. (8.4)

PROOF. Since S normalizes the Fitting subgroup $F(G)$ of G it normalizes each of the p-Sylow subgroups of $F(G)$ also. By (8.3), such a p-Sylow subgroup of $F(G)$ is trivial if $p \neq 7$, so we must have $F(G) = S \triangleleft G$ since $C_G(F(G)) \leqslant F(G)$ ensures that $F(G)$ is non-trivial. Since $C_G(S) = S$ then $G/S \leqslant Aut(S) \cong \mathbb{Z}_6$. Now the result follows from (8.2).

If G is non-solvable then G is non-abelian simple and $|\ N\ | = 21$. (8.5)

PROOF. Let E be a minimal (non-trivial) normal subgroup of G. By (8.3) we get $7| \ |E|$, so that $S \leqslant E$. If $E = S$ then G is solvable, the converse being true by (8.4).

Suppose $E \neq S$. Since E is the direct product of isomorphic simple groups E must itself be simple and non-abelian. Now by a theorem of BURNSIDE [4, Theorem 7.43], if $N_E(S) = S$ then $E = K{:}S$ for some group K of order coprime to 7. This is impossible by (8.3), so $N_E(S) > S$. After (8.2) we get $N_E(S) = N$ has order 21.

Finally, the Frattini argument [4, Theorem 1.37] yields $G = E\cdot N_G(S) = E\cdot N = E$, so that $G = E$ is non-abelian simple, and we are done.

The order of G is not divisible by 5 . (8.6)

PROOF. If false, G has a Sylow 5-subgroup F of order 5, and $C_G(F) = F$ by (7.1). If $N_G(F) = F$ then $G = K{:}F$ for some group K of order prime to 5 by Burnside's theorem [loc cit], against the simplicity of G (cf. (8.4) and (8.5)). So we have $|N_G(F){:}F| = 2$ or 4 since $N_G(F)/F \leqslant Aut(F) \cong \mathbb{Z}_4$.

Now after (8.5) we have $|G| = 2^a{\cdot}3^b{\cdot}5{\cdot}7$ with $b \geqslant 1$. Moreover by Sylow's theorem applied to both $N_G(S)$ and $N_G(F)$ we have $|G:N_G(F)| \equiv 1 \pmod 5$ and $|G:N_G(S)|=1 \pmod 7$, and in the latter case we even know that $|N_G(S)| = 21$. The only possibilities are the following:

(i) $|N_G(F)| = 10, |G| = 2^6{\cdot}3^2{\cdot}5{\cdot}7$.

(ii) $|N_G(F)| = 20, |G| = 2^3{\cdot}3^2{\cdot}5{\cdot}7$.

We use the equations $|G| = \Sigma a(n)$ and $dim V^G = |G|^{-1}\Sigma a(n)\chi(n) \geqslant 5$, eliminate $a(2)$ from them, and arrive at the inequalities

$$0 \leqslant |G|^{-1}[2a(3)+4a(4)+4a(5)+6a(6)+5a(7)+6a(8)] \leqslant 3+\frac{16}{|G|}. \qquad (*)$$

In case (i) we have $a(5) = 2|G|/5$ and $a(7) = 2|G|/7$, which gives

$$0 \leqslant |G|^{-1}[2a(3)+4a(4)+6a(6)+6a(8)] \leqslant 3+\frac{16}{|G|}-\frac{8}{5} = \frac{10}{7},$$

in particular $0 \leqslant \frac{16}{|G|}-\frac{1}{35}$. Thus $|G| \leqslant 16{\cdot}35$, a contradiction.

In case (ii) we have $a(5) = |G|/5$, $a(7) = 2|G|/7$. Moreover in this case $N_G(F)$ contains an element x of order 4. By (7.1) $C_G(x)$ is a 2-group, of order at most 8 since $|C_G(x)|$ divides $|G|$. So either $|C_G(x)| = 8$ and x is not conjugate to its inverse, or else $C_G(x) = \langle x \rangle$ and x is conjugate to its inverse. In either case we get $a(4) \geqslant \frac{|G|}{4}$. Now (*) yields

$$0 \leqslant |G|^{-1}[2a(3)+6a(b)+6a(8)] \leqslant 3+\frac{16}{|G|}-\frac{4}{5}-\frac{10}{7}-1,$$

yielding $|G| \leqslant 70$, contradiction.

Proof of (8.1). We may assume G non-solvable by (8.4), whence it is simple of order $2^a{\cdot}3{\cdot}7$ by (8.6) and (8.5). In fact (8.5) and Sylow's theorem force $|G| = 2^a{\cdot}3{\cdot}7$ with $a = 3$ or 6. If $a = 3$ then $G \cong L_2(7)$, being a simple group of order 168.

It remains to show that $a = 6$ is impossible. In fact we will show that G has a subgroup of index 7, which clearly suffices. Now by a theorem of Frobenius [4, Theorem 7.4.5(a)] there is a 2-group U in G such that $N_G(U)/C_G(U)$ is *not* a 2-group. Choose U with $|U|$ maximal. By (8.3), $N_G(U)$ has order $2^b{\cdot}3$ for some b. Let T be a 2-Sylow of G containing U with $T_1 = N_T(U)$ a 2-Sylow of $N_G(U)$. If $T_1 = T$ then $N_G(U)$ has index 7 in G as required, so we can assume that $T_1 \underset{\neq}{<} T$. Hence $|U| \leqslant 16$ since $|T| = 2^6$. Note that $C_G(U) \leqslant U = O_2(N_G(U))$ by choice of U. Let $R \cong \mathbb{Z}_3$ be a 3-Sylow of $N_G(U)$.

Suppose first that $|U| \leqslant 4$. Then R acts faithfully on U, so $U \cong E_4$ and $U = C_G(U)$. Now from (6.2), we see that T (having index 2 in a 2-Sylow of M_{23}) certainly has a normal subgroup E isomorphic to E_8. Then one checks that $|C_{EU}(U)| \geqslant 8$, against $U = C_G(U)$. So $|U| \geqslant 8$.

Next, since $C_G(U) \leqslant U$ then certainly $Z(T) \leqslant Z(U)$. So if R centralizes $Z(U)$

then $N(Z(T))$ contains both T and R and hence has index 7 in G since $Z(T)$ cannot be normal in G by simplicity. So R does not centralize $Z(U)$, and this is enough to ensure that U is abelian.

If $U \cong E_{16}$ then $U \not\trianglelefteq T$ since all E_{16}'s in a 2-Sylow of M_{23} are normal. This is false, so if $|U| = 16$ then $U \cong \mathbf{Z}_4 \times \mathbf{Z}_4$ by (6.2). Then U is the unique subgroup of T_1 of type $\mathbf{Z}_4 \times \mathbf{Z}_4$, so U is characteristic in T_1, whence normal in T. So in fact $|U| = 8$ and since R does not centralize U then $U \cong E_8$. Let $U_0 = C_U(R) \cong \mathbf{Z}_2$, so that $U_0 = Z(N_G(U))$. By choice of U, $N_G(U_0) = N_G(U)$.

Now $|T_1 : U| = 2$ and U is not the only E_8-subgroup of T_1. Thus T_1 has exactly two E_8-subgroups, call them U and U_1, and $T_1 = UU_1$. This forces $T_2 = N_T(T_1)$ to satisfy $|T_2 : T_1| = 2$. Then U, U_1 are *not* the only two E_8-subgroups of T_2 (otherwise $N_T(U) \not\leq T_1$), so there is $x \in T_2 \setminus T_1$ lying in an E_8-subgroup of T_2. Since $U^x = U_1$, the only possibility is that x centralizes $U \cap U_1 (\cong E_4)$, in particular $x \in C_G(U_0) = N_G(U)$, so $x \in T_1$. This is not the case and (8.1) is proved.

9. The case where $|G|$ is divisible by 5

Here we prove

If 5 divides $|G|$ then G is isomorphic to a subgroup of one of the groups: Σ_5, A_6, $E_{16}:A_5$. (9.1)

We fix a Sylow 5-subgroup F of G. By (7.1) we have $C_G(F) = F$. Let $N = N_G(F)$. Then $N/F \leqslant Aut(F) \cong \mathbf{Z}_4$.

If A is a non-trivial F-invariant subgroup of G of order prime to 5 then $A \cong E_{16}$. (9.2)

Proof. By [4, Theorem 6.2.2] F normalizes a p-Sylow subgroup of A for each prime p; call such an A-invariant p-Sylow S_p. By (7.1) we get $5 \mid |S_p| - 1$, which forces $A = S_2$. Similarly as $C_A(F) = 1$ by (7.1) we must have $|A| = 16$ since $|A| \leqslant 2^7$, and the result follows easily.

The conclusions of (9.1) *hold if G is solvable*. (9.3)

Proof. If $F \leqslant F(G)$ then $F = F(G)$ by (7.1) and the nilpotence of $F(G)$. Then $G/F \leqslant Aut(F) \cong \mathbf{Z}_4$, so $G \leqslant \mathbf{Z}_5 \cdot \mathbf{Z}_4 \leqslant \Sigma_5$.

If $F \not\leqslant F(G)$ then $F(G) \cong E_{16}$ by (9.2). Moreover $F \cdot F(G) \trianglelefteq G$, in fact $F \cdot F(G)/F(G) = F(G/F(G))$, so $G \leqslant E_{16} \cdot (\mathbf{Z}_5 \cdot \mathbf{Z}_4)$. But this latter group contains, besides the identity, 35 involutions, 5·28 elements of order 4, 5·16 elements of order 8 and 2^6 elements of order 5. This yields

$$dim V^G = \frac{1}{320}(24 + 8 \cdot 35 + 4 \cdot 5 \cdot 28 + 2 \cdot 5 \cdot 16 + 4 \cdot 2^6) = 4,$$

contradiction. So in fact $G \leqslant E_{16} \cdot D_{10} \leqslant E_{16} \cdot A_5$, as required.

PROOF OF (9.1). We may assume that G is non-solvable by (9.3). Let E be a minimal normal subgroup of G, and assume first that $5||E|$. If E is solvable then E is an elementary abelian p-group for some p, whence $E = F$ and G is solvable. So E is non-solvable, hence non-abelian simple since it must be a direct product of isomorphic simple groups.

By Brauer's result [2] we get $E \cong A_5$ or A_6. In the first case $G \leqslant Aut(E) \cong \Sigma_5$ and we are done. In the second case, $G/E \leqslant Out(A_6) \cong E_4$, and it is well known that the three subgroups of index 2 in $Aut(A_6)$ are of type $PGL_2(9)$, Σ_6 and M_{10}, respectively (M_{10} is the stabilizer of two points in the action of M_{12} on 12 points). Now $PGL_2(9)$ contains an element of order 10, which shows that G is neither $PGL_2(9)$ nor the full automorphism group $Aut(A_6)$.

The enumeration of elements of Σ_6 is well known, and leads to $dim V^G = 4$. Similarly, a subgroup M_{10} in M_{23} has only four orbits on Ω as is readily checked, and hence, since we know that the character of any M_{10} on $V = H^*(X, \mathbf{Q})$ is the same as that of an M_{10} on $P = \mathbf{Q}\Omega$, we must get $dim V^G = 4$ in this case, too.

So the result is proved if $5||E|$. Assume therefore that 5 does not divide $|E|$. By (9.1), $E \cong E_{16}$ and $E = C_G(E)$ by (7.6). As G is non-solvable then so is G/E and hence a minimal normal subgroup of G/E must be isomorphic to A_5 or A_6. As we have already seen that the group $E_{16}\cdot(\mathbf{Z}_5\cdot\mathbf{Z}_4)$ cannot occur, it follows that $G/E \cong A_5$ or A_6. Now the group M_{23} contains a subgroup $E_{16}\cdot A_6$ with 4 orbits on Ω, and as we have seen before this forces $dim V^G = 4$ if $G \cong E_{16}\cdot A_6$. So in fact $G \cong E_{16}\cdot A_5$.

There are two possibilities for the isomorphism type of E considered as an F_2A_5-module. In the first, A_5 is transitive on the non-identity elements of E; in this case G has a subgroup H of index 5, order $2^6\cdot 3$, with a normal 2-Sylow U and Sylow 3-subgroup R ($\cong Z_3$) satisfying $C_U(R) = 1$. Then one finds (see (10.3)) that $U \cong J(T)$ in the notation of (6.2), in particular E has a complement in U, hence in G by a well known result of GASCHÜTZ ([6, I.17.4]). So $G \cong E_{16} \cdot A_5$ is a split extension isomorphic to the group M_{20} of [3], and certainly lies in M_{23}.

The other possibility is that there are two orbits in the action on $E^\#$, of lengths 5 and 10. In this case the subgroup H of G or order $2^6\cdot 3$ has a normal 2-Sylow U, 3-Sylow R, and $C_E(R) = C_U(R) \cong E_4$. Then H contains 27 involutions, 32 elements of order 3, 36 of order 4, and 96 of order 6, yielding

$$dim V^H = \frac{1}{192}(24+8\cdot 27+6\cdot 32+4\cdot 36+2\cdot 96) = 4.$$

This contradiction completes the proof of (9.1).

10. THE CASE WHERE $|G|$ IS DIVISIBLE BY 9

We turn to the cases in which neither 5 nor 7 divide $|G|$. So G is solvable of order $2^a\cdot 3^b$ with $a \leqslant 7$ and $b \leqslant 2$. Here we prove

If $9 \mid |G|$ then G is isomorphic to a subgroup of one of the following: (10.1)
$E_9 : Q_8$, $E_9 \cdot D_8$, $(A_4 \times A_4) \cdot \mathbf{Z}_2$.

PROOF. Let R be a Sylow 3-subgroup of G, so that $R \cong \mathbf{Z}_3 \times \mathbf{Z}_3$ by (7.1). Suppose first that $R \leqslant F(G)$. By (7.6) we get $R = F(G)$, so as $R = C_G(R)$ then $G/R \leqslant Aut(R) \cong GL_2(3)$. As G/R is a 2-group it must be isomorphic to a 2-subgroup of $GL_2(3)$. This latter group has three maximal subgroups, of type $\mathbf{Z}_8$, Q_8 and D_8 respectively. Now the group $E_9 : \mathbf{Z}_8$ has 9 involutions, 18 elements of order 4, 36 of order 8, and 8 of order 3, yielding

$$dim V^G = \frac{1}{72}(24+8 \cdot 9+4 \cdot 18+2 \cdot 36+6 \cdot 8) = 4.$$

This shows that a 2-Sylow of G cannot contain $\mathbf{Z}_8$, hence lies in Q_8 or D_8, giving the first two possibilities.

Before continuing the proof of (10.1) we interpolate two useful results.

Suppose B is a 2-group with $|B:Z(B)| = 4$. Then its derived group B' (10.2)
has order $|B'| = 2$.

PROOF. Let $|B| = 2^n$. If $x \in B \setminus Z(B)$ then $C_B(x) = Z(B)\langle x \rangle$ has index 2 in B. Thus B has exactly $2^{n-2}+\frac{1}{2}(2^n-2^{n-2})$ conjugacy classes, that is $5 \cdot 2^{n-3}$ classes. On the other hand let $|B'| = 2^c$. Note that $c \geqslant 1$ as B is non-abelian. Then B has $|B:B'| = 2^{n-c}$ characters of degree 1, and since the number of irreducible characters equals the number of conjugacy classes, B has $5 \cdot 2^{n-3} - 2^{n-c}$ irreducible characters of degree $\geqslant 2$. Since $|G|$ equals the sum of the squares of the degrees of the irreducible characters we must have

$$2^n \geqslant 2^{n-c} + 4(5 \cdot 2^{n-3} - 2^{n-c}).$$

This reduces to $2^{n-c} \geqslant 2^{n-1}$, whence $c = 1$ as required.

Suppose B is a 2-group in M_{23} such that $|B| \leqslant 2^6$ and B has an auto- (10.3)
morphism α of order 3 satisfying $C_B(\alpha) = 1$. Then either B is abelian or else $B \cong J(T)$ in the notation of (6.2).

PROOF. Suppose that B is non-abelian. Set $Z = Z(B) \underset{\neq}{<} B$. Since α fixes no elements of $B^\#$ the same is also true of Z and B/Z. If $|B:Z| = 4$ then $|B'| = 2$ by (10.2), so α centralizes B'. This is false, so we have $|B/Z| = 16$ and $|Z|=4$. In fact, $C_z(\alpha)=1$ forces $Z \cong E_4$. Again (10.2) together with $C_{B|Z}(\alpha)=1$ force B/Z abelian, so $B/Z \cong E_{16}$ or $\mathbf{Z}_4 \times \mathbf{Z}_4$. In the latter case, if we choose $x \in B$ so that xZ has order 4 in B/Z, then $\langle x,Z \rangle \cong \mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_4$ or $\mathbf{Z}_2 \times \mathbf{Z}_8$, and both are impossible by (6.2). So in fact $B/Z \cong E_{16}$ and B/Z is generated by subgroups B_1/Z of order 4 which are α-invariant. Now (10.2) yields that each B_1 is abelian, and the result follows from (6.2).

We return to the proof of (10.1) and consider next the case that

$R \cap F(G) = 1$. Then $F = F(G)$ is a 2-group. Now we have $F = <C_F(R_1) | \mathbb{Z}_3 \cong R_1 \leqslant R>$ by [4, Theorem 6.2.4], and since $C_F(R) = 1$ by (7.6) then for each $R_1 \leqslant R$ of order 3 the group $C_F(R_1)$ satisfies the conclusions of (10.3). It follows from (7.6) that $C_F(R_1) = 1$ or E_4 for each such R_1. As $|F| \leqslant 2^6$ there is some $\mathbb{Z}_3 \cong R_0 \leqslant R$ with $C_F(R_0) = 1$, so again (10.3) yields $F \cong E_{16}$ or $F \cong J(T)$ with the notation of (6.2). In the latter case we see that FR contains 27 involutions, $2^7 + 6 \cdot 2^4$ elements of order 3, 36 of order 4, $18 \cdot 2^4$ of order 6, and hence that $dim V^{FR} = 4$. This is impossible, so that $F \cong E_{16}$ and $FR \cong A_4 \times A_4$.

Finally, G/FR must be a subgroup of D_8. If G/FR has an element of order 4 then G has a subgroup H of the form $E_{16} \cdot (E_9 \cdot \mathbb{Z}_4)$. This group contains 51 involutions, 80 elements of order 3, 9·28 of order 4, 9·16 of order 8, and 48 of order 6; we compute that $dim V^H = 4$, contradiction. So G/FR has exponent 2. Let x be an involution of $G \setminus FR$. Now FR has just two normal subgroups isomorphic to A_4, so x either fixes or interchanges them.

If x interchanges these two A_4's then the group $H = FR<x>$ contains 27 involutions, 80 elements of order 3, 36 of order 4 and 144 of order 6; this leads to the contradiction $dim V^{FR<x>} = 4$. So x normalizes each of the A_4's normal in FR. If A_1 is one of these normal A_4's then $A_1 <x> \cong \Sigma_4$ or $\mathbb{Z}_2 \times A_4$. If the second case holds, and if A_2 is the other normal A_4, we must have $A_2 <x> \cong \Sigma_4$ or $\mathbb{Z}_2 \times A_4$, and then $FR<x>$ contains either $\mathbb{Z}_4 \times \mathbb{Z}_3$ or E_{32}, respectively. This contradicts (7.6), so in fact $A_i <x> \cong \Sigma_4$ for $i = 1, 2$, which gives the third possibility of (10.1).

It remains to consider the possibility that $R \cap F(G) \neq 1$, $R \leqslant F(G)$. Then $R \cap F(G) = R_0 \cong \mathbb{Z}_3$, and since $C_G(R) = R$ by (7.6) we get $F(G) \cong \mathbb{Z}_3 \times E_4$ by (7.6) once more. As there are no elements of order 12, another application of (7.6) shows that $C_G(R_0) = F(G)R \cong \mathbb{Z}_3 \times A_4$ and of course $|G:C_G(R_0)| \leqslant 2$. If $G = C_G(R_0)$ then $G \leqslant A_4 \times A_4$. If $|G:C_G(R_0)| = 2$ then $|N_G(R):R| = 2$ and there is an involution $x \in N_G(R)$. We easily see that $G = F(G)R<x>$ is isomorphic either to a subgroup of the group $(A_4 \times A_4) \cdot \mathbb{Z}_2$ already considered, or to the group $\Sigma_3 \times E_4$. But if $G \cong \Sigma_3 \times E_4$ we easily find that $dim V^G = 7$, whereas for the group $K \cong \mathbb{Z}_3 \times E_4$ of index 2 we get $dim V^K = 6$. This contradiction completes the proof of (10.1).

11. End of proof

In this last section we assume that $|G| = 2^a \cdot 3^b$ with $b \leqslant 1$. After (7.3) we may assume that $b = 1$. Note that the last paragraph established

$$G \ncong \Sigma_3 \times E_4 . \tag{11.1}$$

We must show, of course, that G is isomorphic to a subgroup of one of the groups (i)-(xi) in Section 4.

Let $R \cong \mathbb{Z}_3$ be a Sylow 3-subgroup of G, and assume first that $R \leqslant F(G)$, that is $R \trianglelefteq G$. After (7.6) we get $C_G(R) = R \times E$ where E is one of 1, $\mathbb{Z}_2$, or E_4, and of course $|G:C_G(R)| \leqslant 2$. If $E = 1$ then $G \cong \mathbb{Z}_3$ or Σ_3 and we are done. If $E \cong \mathbb{Z}_2$ then either $G = C_G(R) \cong \mathbb{Z}_6$ or else a 2-Sylow of G has order 4 and hence is E_4 or $\mathbb{Z}_4$. If E_4 then $G \cong \mathbb{Z}_2 \times \Sigma_3$ is contained (for example) in Σ_5. If $\mathbb{Z}_4$ then $G \cong \mathbb{Z}_3 \cdot \mathbb{Z}_4$ is contained in the group $(A_4 \times A_4) \cdot \mathbb{Z}_2$ of (vii).

Finally, assume that $E \cong E_4$. If $G = C_G(R)$ then $G \cong \mathbb{Z}_3 \times E_4$ is contained in $A_4 \times A_4$. Otherwise, a Sylow 2-subgroup T of G has order 8. If T is non-abelian then $T \cong D_8$ and G is again contained in $(A_4 \times A_4) \cdot \mathbb{Z}_2$. If T is isomorphic to E_8 then $G \cong \Sigma_3 \times E_4$, against (11.1). The only other possibility is $T \cong \mathbb{Z}_2 \times \mathbb{Z}_4$, and we show this to be impossible as follows: let x be the unique non-identity square in T. Then $G = C_G(x)$, so $G \leqslant \hat{A}_8$ by (7.2), so $G/\langle x \rangle (\cong \mathbb{Z}_2 \times \Sigma_3) \leqslant A_8$. Referring to the proof of (7.2), we see that, up to conjugacy, $G/\langle x \rangle$ must be the subgroup of A_8 given by $\langle(12)(34)(56)(78)\rangle \times \langle(135)(246),(35)(46)\rangle$. In particular, a 2-Sylow of $G/\langle x \rangle$ has the non-identity elements (12)(34)(56)(78), (12)(36)(45)(78) and (35)(46). Pulling these back to G, the first two pull back to E_4's, the third to $\mathbb{Z}_4$. So a 2-Sylow of G, of order 8, contains $\mathbb{Z}_4$ and two distinct E_4's, hence must be D_8. Thus it is not $\mathbb{Z}_2 \times \mathbb{Z}_4$, which is the desired contradiction.

We may now assume that $R \not\leqslant F(G)$, in which case $F(G) = Q$ is a 2-group satisfying $C_G(Q) \leqslant Q$. We set $Y = \Omega_1(Z(Q))$, the subgroup of $Z(Q)$ $(\neq 1)$ generated by its involutions. Suppose that $|Y| \geqslant 4$. Then if Q/Y has an element xY of order $\geqslant 4$ then the abelian group $\langle x, Y \rangle$ contains either $\mathbb{Z}_4 \times E_4$ or $\mathbb{Z}_8 \times \mathbb{Z}_2$. This contradicts (7.6), so Q/Y has exponent 2, that is, it is elementary abelian. Thus the Frattini subgroup $\Phi(Q)$ of Q lies in Y. Now in fact this argument shows that $\Phi(Q) \leqslant Y_0$ whenever $E_4 \cong Y_0 \leqslant Y$. Then we get $\Phi(Q) \leqslant \cap Y_0$ where the intersection runs over the E_4-subgroups of Y. If $|Y| \geqslant 8$ this intersection is trivial, so that $\Phi(Q) = 1$ and Q is itself elementary abelian.

If $|Y| = 4$ then $Z(Q)$ has rank 2. As it admits R, (7.6) yields either $Y = Z(Q) \cong E_4$ or else $Z(Q) \cong \mathbb{Z}_4 \times \mathbb{Z}_4$. In the latter case $\langle x, Z(Q) \rangle$ is abelian of order $\geqslant 2^5$ whenever $x \in Q \setminus Z(Q)$, against (7.6). So if $Z(Q) \cong \mathbb{Z}_4 \times \mathbb{Z}_4$ then $Q = Z(Q)$. Finally, if $|Y| = 2$ then $Z(Q)$ is cyclic, hence centralized by R, so $Y = Z(Q) \cong \mathbb{Z}_2$ by (7.6) once more. So there are the following possibilities for Q:

(a) $Q \cong E_{2^a}$, $2 \leqslant a \leqslant 4$.

(b) $Q \cong \mathbb{Z}_4 \times \mathbb{Z}_4$.

(c) $Z(Q) \cong E_4$, $Q \neq Z(Q)$.

(d) $Z(Q) \cong \mathbb{Z}_2$.

If (b) *holds then* $G \leqslant E_{16} : \Sigma_4$ *(group* (x)*).* (11.2)

PROOF. The group $L = E_{16} : \Sigma_4$ is the (unique) extension indicated which has no elements of order 6. By (10.3) we see that $O_2(L) \cong J(T)$ in the notation of (6.2). Moreover $L \cong J(T) : \Sigma_3$. Now a 3-Sylow P of L fixes each of the three $\mathbb{Z}_4 \times \mathbb{Z}_4$-subgroups of $J(T)$ (cf. 6.2)), so at least one of these is fixed by $N_L(P) \cong \Sigma_3$, yielding a subgroup $(\mathbb{Z}_4 \times \mathbb{Z}_4) : \Sigma_3$ within L.

Finally, since $G \cong (\mathbb{Z}_4 \times \mathbb{Z}_4){:}\mathbb{Z}_3$ or $(\mathbb{Z}_4 \times \mathbb{Z}_4){:}\Sigma_3$ the result follows.

If (c) *holds then* $C_Q(R) \cap Z(Q) = 1$. (11.3)

PROOF. If not then R centralizes $Z(Q) \cong E_4$, so that $Y = Z(Q) = C_Q(R)$ by (7.6). Now we have already seen that Q/Y is elementary abelian, and since $C_{Q/Y}(R) = 1$ then $Q/Y \cong E_4$ or E_{16}. In the latter case there are 5 distinct R-invariant subgroups D/Y of Q/Y of order 16 which partition Q/Y. Since Q is not elementary abelian, some D is not E_{16}, so as $C_D(R) = Y$ then D cannot even be abelian.

So whatever the possibility for Q/Y, there is a non-abelian subgroup of D of order 16 which admits R and $C_D(R) = Z(D) \cong E_4$. Using (10.2) we see that $D \cong \mathbb{Z}_2 \times Q_8$, $DR \cong \mathbb{Z}_2 \times SL_2(3)$. But then DR has 3 involutions, 8 elements of order 3, 12 of order 4, and 24 of order 6. This yields the contradiction

$$dim V^{DR} = \frac{1}{48}(24+8{\cdot}3+6{\cdot}8+4{\cdot}12+2{\cdot}24) = 4.$$

If (c) *holds then either* $|Q| = 2^6$, $C_Q(R) = 1$ *and* $G \leqslant E_{16} : \Sigma_4$ *(type* (x)); *or* $|Q| = 2^5$, $C_Q(R) \cong \mathbb{Z}_2$ *and* $G \leqslant E_{16} : D_{12}$ *(type* (viii)). (11.4)

PROOF. We have $Y = Z(Q) \cong E_4$, Q/Y elementary abelian, and $C_Y(R) = 1$. If $|Q| = 16$ then $|Q'| = 2$ by (10.2), so $1 \neq Q' \leqslant Y \cap C(R)$, contradiction. So $|Q| \geqslant 2^5$. We have $Q/Y = B/Y \times D/Y$ where $B/Y = C_{Q/Y}(R)$, and $|B{:}Y| \leqslant 4$ by (7.6). If $B/Y = 1$ we may apply (10.3) to see that $Q = D \cong J(T)$ in the notation of (6.2). If now $G = QR$ then clearly $G \leqslant E_{16} : \Sigma_4$, in fact $G \cong E_{16}{:}A_4$, the subgroup of index 2. If $G \neq QR$ then a 2-Sylow U of G has order 2^7 and $G \cong Q{:}\Sigma_3$. In this case each of the two E_{16}'s in Q are normal in U by (6.2) (since U is a 2-Sylow of M_{23}), hence again $G \cong E_{16}{:}\Sigma_4$.

Now assume that $B/Y \neq 1$. Since $C_D(R) = 1$ we can apply (10.3) to D and conclude that either D is abelian of order 16 or $D \cong J(T)$ in the notation of (6.2). Assume first that $D \cong E_{16}$. If $|B{:}Y| = 4$ then $B = Y \times C_B(R) \cong E_{16}$ (by (7.6)), and the group QR is isomorphic to the group denoted by H in the last paragraph of the proof of (9.1). That calculation therefore gives $dim V^{QR} = 4$, a contradiction. So if $D \cong E_{16}$ then $|B{:}Y| = 2$, that is $C_Q(R) \cong \mathbb{Z}_2$. Thus $QR = D{\cdot}R{\cdot}C_Q(R) \cong E_{16}{:}\mathbb{Z}_6$. If $QR = G$ we are done. If there is an involution of G inverting R then $G \cong E_{16}{:}D_{12}$ as required. Otherwise $G \cong E_{16}{:}(\mathbb{Z}_3{\cdot}\mathbb{Z}_4)$, G contains 19 involutions, 32 elements of order 3, 60 of order 4, 32 of order 6, and 48 of order 8, yielding the contradiction

$$dim V^{G} = \frac{1}{192}(24+8{\cdot}19+6{\cdot}32+4{\cdot}60+2{\cdot}32+2{\cdot}48) = 4.$$

So we may now assume $B/Y \neq 1$ and $D \cong E_{16}$. So either $D \ncong \mathbb{Z}_4 \times \mathbb{Z}_4$ or else $D \cong J(T)$ and D contains an R-invariant subgroup $D_0 \cong \mathbf{Z}_4 \times \mathbf{Z}_4$. Choose an involution $x \in C_B(R)$ and let $H = DR<x>$ or $D_0R<x>$ in the two cases. Consider $O_2(H)$: if all involutions of $O_2(H)$ lie in $Y<x> \cong E_8$ then $O_2(H)$ contains 7 involutions and 24 elements of order 4, whence $dim V^{O_2(H)} = \frac{1}{3}(24+8\cdot 7+4\cdot 24)$ is not an integer. So there are involutions in $O_2(H)-Y<x>$. These being permuted by R in cycles of length 3, we see that $O_2(H)$ contains 19 involutions and 12 elements of order 4. So in this case $dim V^{O_2(H)} = 1/32(24+8\cdot 19+4\cdot 12) = 7$. On the other hand $D_1 \cong \mathbf{Z}_4 \times \mathbf{Z}_4$ satisfies $dim V^{D_1} = \frac{1}{16}(24+8\cdot 3+4\cdot 12) = 6$, yielding the contradiction $D_1 \leqslant O_2(H)$, $dim V^{D_1} < dim V^{O_2(H)}$. This completes the proof of (11.4).

If (d) *holds then either* $G \leqslant GL_2(3)$ *(type* (xi)) *or* $G \cong (Q_8 \star Q_8):\mathbf{Z}_3$ *(type* (11.5) (ix)).

PROOF. Let $Y = Z(Q) \cong \mathbf{Z}_2$. If $|Q| = 8$ we must have $Q \cong Q_8$ and $QR \cong SL_2(3)$. If $G = QR$ we are done and if there is an involution in $G \setminus QR$ then $G \cong GL_2(3)$. Otherwise, a 2-Sylow of G is isomorphic to Q_{16}, G has 1 involution, 8 elements of order 3, 18 of order 4, 8 of order 6, and 12 of order 8, yielding

$$dim V^G = \frac{1}{48}(24+8+6\cdot 8+4\cdot 18+2\cdot 8+2\cdot 12) = 4.$$

So we may now assume that $|Q| \geqslant 16$.

Note that by (7.2), we get $G/Y \leqslant A_8$, in particular $|Q/Y| \leqslant 2^5$. If Q/Y is elementary abelian then $\Phi(Q) = Z(Q) = Y$, so $Q \cong Q_8 \star Q_8$ (cf. [4, Theorem 5.4.9]) and $QR \cong (Q_8 \star Q_8):\mathbf{Z}_3$. Thus if $G = QR$ we are done. If not, we see that, up to conjugacy, G/Y is the group $<(12)(34),(13)(24),(56)(78),(57)(68),(123)(456),(23)(56)>$. Then using (3.4), we see that G itself contains 19 involutions, 32 elements of order 3, 60 of order 4, 32 of order 6, and 48 of order 8. This gives

$$dim V^G = \frac{1}{192}(24+8\cdot 19+6\cdot 32+4\cdot 60+2\cdot 32+2\cdot 48) = 4.$$

Now assume that Q/Y is not elementary abelian. If $C_{Q/Y}(R) = 1$ then $|Q:Y| = 16$ and we get $Q/Y \cong \mathbf{Z}_4 \times \mathbf{Z}_4$ by (10.2). As A_8 has no subgroup of the shape $(\mathbf{Z}_4 \times \mathbf{Z}_4):\mathbf{Z}_3$ this is impossible, so $C_{Q/Y}(R) \cong \mathbf{Z}_2$ and $C_Q(R) \cong E_4$. Setting $B = C_Q(R)$, since $N_Q(B) > B$ and R acts on $N_Q(B)/B$ without non-trivial fixed-points, it must be the case that $C_Q(B) = N_Q(B)$ has order 2^4 or 2^6. Since $|Q| \leqslant 2^6$ and $Z(Q) \cong \mathbf{Z}_2$ we must have $|C_Q(B)| = 2^4$ and $|Q| = 2^6$. Looking at the subgroups of A_8 ($\geqslant QR/Y$), we see that Q/Y has a subgroup $D/Y \cong E_{16}$ which admits R with $C_D(R) = Y$. Then $D \cong Q_8 \star Q_8$, $QR \cong (Q_8 \star Q_8):Z_6$ and QR contains 27 involutions, 32 elements of order 3, 36 of order 4, and 96 of order 6. This yields

$$dim V^{QR} = \frac{1}{192}(24+8 \cdot 27+6 \cdot 32+4 \cdot 36+2 \cdot 96) = 4,$$

a contradiction. This completes the proof of (11.5).

It remains to consider case (a), i.e., Q elementary abelian. This is easy. If $Q \cong E_4$ then $G \cong A_4$ or Σ_4. If $Q \cong E_8$ then $QR \cong \mathbb{Z}_2 \times A_4$. If $G = QR$ or if there is an involution inverting R then $G \leqslant \mathbb{Z}_2 \times \Sigma_4$, which is contained in the group $(A_4 \times A_4):\mathbb{Z}_2$ (type (vii)). If there is no involution inverting R and $G \neq QR$ then $G \cong E_4 \cdot (\mathbb{Z}_3 \cdot \mathbb{Z}_4)$ is still isomorphic to a subgroup of $(A_4 \times A_4):\mathbb{Z}_2$ (type (vii)).

Finally assume that $Q \cong E_{16}$. If $C_Q(R) = 1$ then $G \leqslant E_{16}:\Sigma_3 \leqslant E_{16}:\Sigma_4$ (type (x)). If $C_Q(R) \neq 1$ then $C_Q(R) \cong E_4$. If $G = QR$ then $G \cong E_4 \times A_4 \leqslant (A_4 \times A_4):\mathbb{Z}_2$ (type (vii)), and if there is an involution x inverting R then x cannot centralize $C_Q(R)$ by (7.6) (for in this case $Q<x>$ contains $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_4$). Thus $C_Q(R)<x> \cong D_8$ and again $G \leqslant (A_4 \times A_4):\mathbb{Z}_2$ (type (vii)). The last possibility is that no involution of G inverts R. In this case the group $N_G(R)$ has 2-Sylow $\mathbb{Z}_2 \times \mathbb{Z}_4$, and we showed earlier (when considering the case $R \leqslant F(G)$) that G cannot contain such a group. This completes the proof of the Main Theorem.

REFERENCES

1. W. BARTH, C. PETERS, A. VAN DE VEN (1984). *Compact Complex Surfaces, Ergebnisse, 3 Folge, Band 4,* Springer.
2. R. BRAUER (1980). *Simple Groups of Order* $5 \cdot 2^a \cdot 3^b$, *Collected works, vol. II,* 421-470, MIT Press.
3. J. CONWAY (1971). Three lectures on exceptional groups. *Finite Simple Groups,* Academic Press.
4. D. GORENSTEIN (1968). *Finite Groups,* Harper and Row.
5. D. GORENSTEIN, K. HARADA (1974). *Finite Groups whose 2-Subgroups are Generated by at most 4 Elements,* Mem. AMS. No. 147.
6. B. HUPPERT (1967). *Endliche Gruppen I, Grundlehren, Band 134,* Springer.
7. S. MUKAI, *Finite Groups of Automorphisms of K3-Surfaces and the Mathieu group,* preprint.
8. V.V. NIKULIN (1980). Finite groups of automorphisms of Kählerian surfaces of type $K3$. English translation in *Moscow Math. Soc. 30,* 71-137.

Ternary Continued Fractions and the Evenly-Tempered Musical Scale

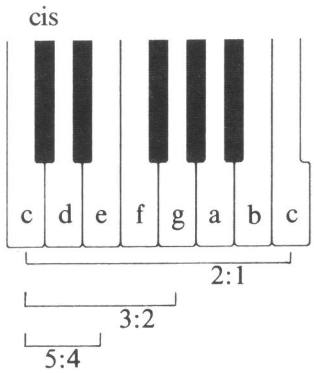
John T. Kent
Department of Statistics, University of Leeds
Leeds, LS2 9JT, U.K.

In the evenly-tempered 12-note chromatic scale of western music, two important intervals are well-approximated - the pure fifth with ratio 3/2, and the major third with ratio 5/4. After taking log-ratios, the musical scale can be viewed as an example of the approximation of two irrational numbers by a pair of rational numbers with the same denominator (in this case, 12). A general approach to such problems is provided by the theory of ternary continued fractions.

1. Introduction

The most important intervals in music are the octave, pure fifth, and major third. Perhaps they are consonant to the ear because they are based on simple whole number ratios (2:1, 3:2, and 5:4, respectively). These intervals arise naturally out of the overtone series for a vibrating string. A string vibrating at a frequency f also vibrates at $2f$, $3f$, $4f$, etc. The ratios between the overtones include these basic intervals. There are more complicated intervals than the major third, but they are not heard by the ear strongly enough to be a major factor in tuning.

It is impossible to tune a scale so that all of these intervals come out exact for all the notes. In every system of tuning the ratio of two notes an octave apart is always taken to be exactly 2:1. The attempt of most systems of tuning has been to approximate the fifth and major third, although the accuracy of the fifth often predominates.



For example, the Pythagorean system of tuning is based only on the octave and fifth. All the fifths but one have a ratio of 3:2. It works well for unison melodies and simple melodies with fourths and fifths, but for more complicated melodies with thirds and sixths it sounds dissonant. Furthermore, because some half-steps have different ratios than others, some keys sound more consonant than others. A piece played in C will sound more in tune than a piece in $F\#$.

The problems of different tunings in different keys can be avoided by restricting attention to tunings in which all the half-steps have the same ratio. Such a scale is called evenly-tempered. The modern day piano and many other instruments are tuned to the 12-note evenly-tempered scale. The basic interval is the half-step.

More generally, we might consider an evenly-tempered scale of n notes in which each unit interval has a ratio $2^{1/n}:1$. If there are k notes in the approximation to the pure fifth, then this interval has a ratio of $2^{k/n}:1$. This will be an irrational number and so can never be exactly 3:2. Similarly for the major third. Thus, in a good evenly-tempered n-note scale there will be numbers h and k such that $2^{h/n} \approx 5/4$ and $2^{k/n} \approx 3/2$. That is, $h/n \approx \log_2 5/4$ and $k/n \approx \log_2 3/2$.

In fact, logarithms make intuitive sense in dealing with intervals. The ear hears a half-step above a half-step as a whole step, so it seems more natural to add the logarithms of the ratios than to multiply the ratios. If one divides the octave into 1200 logarithmic units known as cents, each half-step of the 12-note scale is 100 cents. A true major third has a value of $1200(\log_2 5/4) \approx 386$ cents and a true pure fifth has a value of $1200(\log_2 3/2) \approx 702$ cents. Thus in the 12-note scale the major third is 14 cents sharper than a true major third and the fifth is 2 cents flat.

The object of this paper is to find evenly-tempered scales which give good approximations to the true major third and pure fifth. We are looking for a sequence of pairs of rational numbers with the same denominator, which approximate $\log_2 5/4$ and $\log_2 3/2$. The process by which this is done is known as the theory of ternary continued fractions, and it forms an extension of the idea of ordinary continued fractions. Ordinary continued fractions are used to approximate one number; ternary continued fractions are used to approximate two numbers. They were developed by JACOBI [6]. We will discuss Jacobi's expansion in Section 3 and then an alternative expansion known as the reversed expansion in Section 4. These two expansions can be combined (Section 5) in a way to give us a sequence of approximations which include many of the scales proposed by musical theorists (Section 6).

The main ideas of this paper are due to BARBOUR ([2], and [3], Chapter 6), who gives further details about the musical implications of these results. Here we shall concentrate in more detail on the mathematical development of ternary continued fractions.

2. Ordinary continued fractions

The theory of ordinary continued fractions provides a powerful method of finding a sequence of rational approximations to an irrational number; see, for example HARDY and WRIGHT ([5], Chapter 10) or BAKER ([1], Chapter 6).

For our purposes we shall restrict attention to the case where $\alpha_0 < 1$ is an irrational positive number. The continued fraction expansion for α_0 can be developed as follows. Define a sequence of integers $p_i \geqslant 1$ and positive real numbers $\alpha_i < 1$ by

$$p_i = [\alpha_i^{-1}], \quad \alpha_{i+1} = \alpha_i^{-1} - p_i, \quad i \geqslant 0,$$

where [.] denotes integer part. At each stage we approximate a remainder by the integer part of its reciprocal. Thus,

$$\alpha_i = [\alpha_{i+1} + p_i]^{-1}. \tag{2.1}$$

If for some j we approximate $\alpha_{j+1} \approx 0$ and use backwards recursion in (2.1), then we obtain a rational approximation for α_0 which can be written as a ratio of integers, A_j / B_j say.

For more formal mathematical work it is convenient to set out this procedure in terms of linear transformations. Set $U_0 = \alpha_0$, $V_0 = 1$, and define sequences of positive real numbers by

$$\begin{bmatrix} U_{i+1} \\ V_{i+1} \end{bmatrix} = \begin{bmatrix} -p_i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} U_i \\ V_i \end{bmatrix}$$

where $p_i = [V_i / U_i]$. Thus V_i / U_i is the same as α_i^{-1} above. Next define sequences of integers A_i, B_i by

$$\begin{bmatrix} A_i \\ B_i \end{bmatrix} = \begin{bmatrix} A_{i-2} & A_{i-1} \\ B_{i-2} & B_{i-1} \end{bmatrix} \begin{bmatrix} 1 \\ p_i \end{bmatrix}$$

with initial conditions

$$\begin{bmatrix} A_{-2} & A_{-1} \\ B_{-2} & B_{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then A_i / B_i is called the *i-th convergent* for α_0 and $A_i / B_i \rightarrow \alpha_0$ as $i \rightarrow \infty$. This is classically expressed by the equality

$$\alpha_0 = \cfrac{1}{p_0 + \cfrac{1}{p_1 + \cfrac{1}{p_2 + \cdots}}}.$$

Further it can be shown that the speed of convergence is quite rapid,

$$| A_i / B_i - \alpha_0 | < 1 / B_i^2. \tag{2.2}$$

The multivariate version of Dirichlet's theorem (see for example HARDY and WRIGHT [5], pp. 169-170, or BAKER [1], pp. 56-59) says that, given positive numbers $\beta_1, \ldots, \beta_m$ and an integer $Q^* > 0$, there exist integers $Q \leqslant Q^*$ and

$P_1, \ldots, P_m$ such that

$$| P_j/Q - \beta_j | < Q^{-1-1/m}, \quad j = 1, \ldots, m .$$

From (2.2) we see that ordinary continued fractions ($m = 1$) always achieve this inequality. Unfortunately the approximations from ternary continued expansions ($m = 2$) are not so powerful in general.

3. Jacobi's ternary continued fraction expansion

Let $U_0 < V_0 < W_0 = 1$ be three positive numbers. The objective is to find integers (A_i, B_i, C_i) such that $A_i\colon B_i\colon C_i$ approximates $U_0\colon V_0\colon W_0$. To ensure the expansion is well defined, suppose U_0, V_0, W_0 are linearly independent over the rationals.

First define sequences U_i, V_i, W_i by the following recurrence formulae for $i \geqslant 0$,

$$\begin{pmatrix} U_{i+1} \\ V_{i+1} \\ W_{i+1} \end{pmatrix} = \begin{pmatrix} -p_i & 1 & 0 \\ -q_i & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} \tag{3.1}$$

where

$$p_i = [V_i/U_i] \ , \ q_i = [W_i/U_i] .$$

Next define 3 sequences of integers recursively for $i \geqslant 0$ by

$$\begin{pmatrix} A_i \\ B_i \\ C_i \end{pmatrix} = \begin{pmatrix} A_{i-3} & A_{i-2} & A_{i-1} \\ B_{i-3} & B_{i-2} & B_{i-1} \\ C_{i-3} & C_{i-2} & C_{i-1} \end{pmatrix} \begin{pmatrix} 1 \\ p_i \\ q_i \end{pmatrix} \tag{3.2}$$

with initial conditions

$$\begin{pmatrix} A_{-3} & A_{-2} & A_{-1} \\ B_{-3} & B_{-2} & B_{-1} \\ C_{-3} & C_{-2} & C_{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

This is the Jacobi expansion for (U_0, V_0, W_0) (JACOBI [6]); see also DAUS [4]. The triple (A_i, B_i, C_i) is known as the *i-th convergent set*, the pair (q_i, p_i) is the *i-th partial quotient set*, and the triple (U_i, V_i, W_i) is the *i-th complete quotient set*. Thus, the approximation of $U_0\colon V_0\colon W_0$ can be depicted as

$$1 : p_0 + \cfrac{1}{q_1 + \cfrac{p_2 + \cfrac{1}{q_3 + \cdots}}{q_2 + \cfrac{p_3 + \cdots}{q_3 + \cdots}}} : q_0 + \cfrac{p_1 + \cfrac{1}{q_2 + \cfrac{p_3 + \cdots}{q_3 + \cdots}}}{q_1 + \cfrac{p_2 + \cfrac{1}{q_3 + \cdots}}{q_2 + \cfrac{p_3 + \cdots}{q_3 + \cdots}}} .$$

The 3×3-matrix in (3.1) has determinant $+1$, so that $(U_{i+1}, V_{i+1}, W_{i+1})$ will be linearly independent when (U_i, V_i, W_i) are. In particular $U_{i+1}\neq 0$ which is the only requirement that needs to be satisfied in order to continue the expansion for another step. Thus, the expansion can be continued infinitely when (U_0, V_0, W_0) are linearly independent.

From the definition of p_i and q_i we see that $U_{i+1}<U_i$ and $V_{i+1}<U_i$. Thus, since $W_{i+1} = U_i$, we get

$$U_{i+1}<W_{i+1}\,,\quad V_{i+1}<W_{i+1}\,.$$

Therefore, at each step $0\leqslant p_i\leqslant q_i$ and $q_i\geqslant 1$.

Let $\sigma_{1,i}=V_i/U_i$, $\sigma_{2,i}=W_i/U_i$. Then $\sigma_{1,i}<\sigma_{2,i}$ and $\sigma_{2,i}>1$ at each step. Note that $\sigma_{1,i}$ and $\sigma_{2,i}$ play a role similar to that of α_i^{-1} in Section 2. Inverting the matrix in (3.1) we see that

$$\begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & p_i \\ 0 & 1 & q_i \end{pmatrix} \begin{pmatrix} U_{i+1} \\ V_{i+1} \\ W_{i+1} \end{pmatrix} \tag{3.3}$$

Thus

$$\sigma_{1,i} = p_i+\frac{1}{\sigma_{2,i+1}}\,,\quad \sigma_{2,i} = q_i+\frac{\sigma_{1,i+1}}{\sigma_{2,i+1}}\,,$$

where

$$1/\sigma_{2,i+1}<1 \text{ and } \sigma_{1,i+1}/\sigma_{2,i+1}<1\,.$$

Also, if $p_i=q_i$, then, since $\sigma_{1,i}<\sigma_{2,i}$, we must have $\sigma_{1,i+1}>1$, and so $p_{i+1}\geqslant 1$.

It can be shown (Perron [7], [8]) that these properties of the partial quotient sets (that is, that $0\leqslant p_i\leqslant q_i$ and $q_i\geqslant 1$ for all i, and that $p_i=q_i\Rightarrow p_{i+1}\geqslant 1$) guarantee that the expansion is unique.

The first convergent sets are

$$\begin{aligned} A_0 &= 1 & B_0 &= p_0 & C_0 &= q_0 \\ A_1 &= q_1 & B_1 &= q_1p_0+1 & C_1 &= q_1q_0+p_1\,. \end{aligned}$$

Thus

$$A_1>0,\quad B_1>0,\quad C_1>0\,.$$

Let S_i stand for either A_i, B_i, or C_i. Then (3.2) together with the facts that $q_i\geqslant 1$ and $S_i>0$ for $i\geqslant 1$ show that $S_i>S_{i-1}$ for $i\geqslant 4$. Therefore $\{S_i\}_{i\geqslant 3}$ forms a strictly increasing sequence (Daus [4], p. 281).

The following identity is easily proved by induction for $i\geqslant 0$ using (3.3):

$$\begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix} = \begin{pmatrix} A_{i-3} & A_{i-2} & A_{i-1} \\ B_{i-3} & B_{i-2} & B_{i-1} \\ C_{i-3} & C_{i-2} & C_{i-1} \end{pmatrix} \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix}. \tag{3.4}$$

Dividing two rows in (3.4) by U_i and taking their ratio gives

$$\frac{U_0}{W_0} = \frac{A_{i-3}+\sigma_{1,i}\, A_{i-2}+\sigma_{2,i}\, A_{i-1}}{C_{i-3}+\sigma_{1,i}\, C_{i-2}+\sigma_{2,i}\, C_{i-1}} \tag{3.5}$$

If we approximate $\sigma_{1,i}$ by p_i and $\sigma_{2,i}$ by q_i we might hope that we would get an approximation to U_0/W_0. That is, we would hope,

$$\frac{A_{i-3}+p_i\, A_{i-2}+q_i\, A_{i-1}}{C_{i-3}+p_i\, C_{i-2}+q_i\, C_{i-1}} = \frac{A_i}{C_i} \approx \frac{U_0}{W_0}\ .$$

In fact A_i/C_i and B_i/C_i do converge to U_0/W_0 and V_0/W_0 respectively, but convergence is not nearly so swift as it is for ordinary continued fractions. We shall show that

$$|\, A_i/C_i - U_0/W_0 \,| + |\, B_i/C_i - V_0/W_0 \,| = O(C_i^{-1})$$

in contrast to (2.2) for ordinary continued fractions.

THEOREM 1. (VAISÄLÄ [9]). *Let $U_0<V_0<W_0$ be positive numbers linearly independent over the rationals. Let $(A_i,\, B_i,\, C_i)$ be the i-th convergent set in the Jacobi expansion. Then $A_i/C_i \rightarrow U_0/W_0$ and $B_i/C_i \rightarrow V_0/W_0$ as $i \rightarrow \infty$.*

PROOF. Set $H_i = A_i-(U_0/W_0)C_i$ and $K_i = B_i-(V_0/W_0)C_i$. Then

$$A_i/C_i - U_0/W_0 = H_i/C_i\ ,\quad B_i/C_i - V_0/W_0 = K_i/C_i\ .$$

Now $C_i \uparrow \infty$ so if there is an upper bound on H_i and K_i, convergence is guaranteed.

From (3.5) we see that

$$H_{i-3}+\sigma_{1,i}\, H_{i-2}+\sigma_{2,i}\, H_{i-1} = 0\ .$$

Therefore,

$$H_{i-1} = -1/\sigma_{2,i}\, [H_{i-3}+\sigma_{1,i}\, H_{i-2}]\ .$$

Replacing i by $i+1$ we get

$$H_i = -1/\sigma_{2,i+1}\, [H_{i-2}+\sigma_{1,i+1}\, H_{i-1}]\ . \tag{3.6}$$

Also, from (3.2) we have

$$H_i = H_{i-3}+p_i\, H_{i-2}+q_i\, H_{i-1}\ . \tag{3.7}$$

There are two cases to consider. If H_{i-2}, H_{i-1} have the same sign, then by (3.6) H_i has the opposite sign, so from (3.7)

$$|\, H_i \,| = |\, H_{i-3} \,| - p_i\, |\, H_{i-2} \,| - q_i\, |\, H_{i-1} \,| < |\, H_{i-3} \,|\ .$$

If H_{i-2}, H_{i-1} have the opposite signs, then from (3.6)

$$\begin{aligned} |\, H_i \,| &< \max\, \{\sigma_{2,i+1}^{-1}\, |\, H_{i-2} \,|\, ,\, (\sigma_{1,i+1}/\sigma_{2,i+1})\, |\, H_{i-1} \,|\, \} \\ &< \max\, \{\, |\, H_{i-2} \,|\, ,\, |\, H_{i-1} \,|\, \}. \end{aligned}$$

Therefore in either case $| H_i | < \max \{ | H_{i-3} | , | H_{i-2} | , | H_{i-1} | \}$.

Thus $| H_i |$ is bounded above by $\max \{ | H_0 | , | H_1 | , | H_2 | \}$. Similarly reasoning puts an upper bound on $| K_i |$. Hence the theorem follows. □

In the theory of ordinary continued fractions, we know that any infinite sequence of positive integers is the continued fraction expansion of some positive irrational number. We have an analogous result for ternary continued fractions. Let $\{(p_i, q_i)\}_i$ be a sequence of pairs of integers such that $0 \leqslant p_i \leqslant q_i$, $q_i \geqslant 1$ for all i, and if $p_i = q_i$ then $p_{i+1} \geqslant 1$. Then $\{(p_i, q_i)\}_i$ is the Jacobi expansion for some pair of positive numbers α_0 and β_0 and $\{\alpha_0, \beta_0, 1\}$ are linearly independent over the rationals. For further details see PERRON [7], [8].

The methods of this section can be extended in a straightforward way to give a sequence of simultaneous rational approximations to more than two irrational numbers; see PERRON [7].

4. THE REVERSED EXPANSION

There is another expansion due to BARBOUR [2] that can be used to get approximations to $U_0 : V_0 : W_0$. In the Jacobi expansion one always divided by U_i at the i-th step. In the following, the reversed expansion, one divides by V_i. It is defined by the following recursion formulae in matrix form,

$$\begin{pmatrix} U_{i+1} \\ V_{i+1} \\ W_{i+1} \end{pmatrix} = \begin{pmatrix} 1 & -p_i & 0 \\ 0 & -q_i & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} \tag{4.1}$$

where

$$p_i = [U_i / V_i] \text{ and } q_i = [W_i / V_i] .$$

Set

$$\begin{pmatrix} A_i \\ B_i \\ C_i \end{pmatrix} = \begin{pmatrix} 1 & A_{i-2} & A_{i-1} \\ 0 & B_{i-2} & B_{i-1} \\ 0 & C_{i-2} & C_{i-1} \end{pmatrix} \begin{pmatrix} p_i \\ 1 \\ q_i \end{pmatrix} \tag{4.2}$$

with the same initial conditions as the Jacobi expansion (cf. (3.2)). Schematically, we can write for this approximation of $U_0 : V_0 : W_0$

$$p_0 + \cfrac{p_1 + \cfrac{p_2 + \cfrac{p_3 + \cdots}{q_3 + \cdots}}{q_2 + \cfrac{1}{q_3 + \cdots}}}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdots}}} : 1 : q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cdots}}} .$$

It is interesting to note that $\{q_i\}_{i\geq 0}$ is just the ordinary continued fraction expansion for V_0/W_0. Also, the B_i/C_i are just the ordinary convergents. Thus the reversed expansion is not as symmetrical as the Jacobi expansion. □

Since the 3×3-matrix in (4.1) has determinant 1, we have again that, if U_i, V_i, W_i are linearly independent over the rationals, then U_{i+1}, V_{i+1}, W_{i+1} will be, and in particular $V_{i+1}\neq 0$. Thus, if U_0, V_0, W_0 are linearly independent, we can define an infinite reversed expansion for them.

In matrix form we have the following identities for $i\geq 0$,

$$\begin{pmatrix} U_0 \\ V_0 \\ W_0 \end{pmatrix} = \begin{pmatrix} 1 & A_{i-2} & A_{i-1} \\ 0 & B_{i-2} & B_{i-1} \\ 0 & C_{i-2} & C_{i-1} \end{pmatrix} \begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix}$$

which follow easily from (4.1) by induction as

$$\begin{pmatrix} U_i \\ V_i \\ W_i \end{pmatrix} = \begin{pmatrix} 1 & 0 & p_i \\ 0 & 0 & 1 \\ 0 & 1 & q_i \end{pmatrix} \begin{pmatrix} U_{i+1} \\ V_{i+1} \\ W_{i+1} \end{pmatrix} \tag{4.3}$$

Let

$$\tau_{1,i} = U_i/V_i \qquad \tau_{2,i} = W_i/V_i\ .$$

Then the recursion formulae become

$$\tau_{1,i} = p_i + \tau_{1,i+1}/\tau_{2,i+1} \qquad \tau_{2,i} = q_i + 1/\tau_{2,i+1}$$

where

$$\tau_{1,i+1}/\tau_{2,i+1} < 1 \quad \text{and} \quad 1/\tau_{2,i+1} < 1\ .$$

As in the Jacobi expansion we have convergence of A_i: B_i: C_i to U_0: V_0: W_0, but the proof will be postponed until the next section where a more general theorem is proved.

5. The mixed expansion

In the problem of finding evenly-tempered scales one is interested in scales with small numbers of notes. In both the Jacobi and reversed expansions C_i increases too rapidly to give many interesting scales; see Tables 1(a) and 1(b). So the slow mixed expansion was devised by Barbour [2] to slow the growth of C_i. At each step of the slow mixed expansion, one divides by U_i or V_i whichever is larger.

TABLE 1. Ternary continued expansions for ($\log_2 5/4$, $\log_2 3/2$, 1), adapted from BARBOUR [2].

(a) **Jacobi Expansion**

p_i	q_i	A_i	B_i	C_i	**Error** A_i	**Error** B_i	**Total Error**	H_{i+3}	K_{i+3}
1	3	1	1	3	14	—302	316	0.03	—0.75
0	1	1	2	3	14	98	112	0.03	0.24
1	7	8	15	25	— 2	18	20	—0.04	0.37
0	1	9	16	28	— 1	—16	17	—0.01	—0.38
0	1	10	18	31	0.8	— 5.2	6.0	0.02	—0.13
0	2	28	51	87	— 0.1	1.4	1.5	—0.008	0.1

(b) **Reversed Expansion**

p_i	q_i	A_i	B_i	C_i	**Error** A_i	**Error** B_i	**Total Error**	H_{i+3}	K_{i+3}
0	1	0	1	1	—386	498	884	—0.32	0.42
0	1	0	1	2	—386	—102	488	—0.64	—0.17
1	2	1	3	5	—146	18	164	—0.61	—0.08
2	2	4	7	12	14	— 2	16	0.14	—0.02
0	3	13	24	41	— 6	1	7	—0.20	0.02
0	1	17	31	53	— 1.4	— 0.1	1.5	—0.06	—0.003

(c) **Slow Mixed Expansion**

Step	p_i	q_i	A_i	B_i	C_i	**Error** A_i	**Error** B_i	**Total Error**	H_{i+3}	K_{i+3}
R	0	1	0	1	1	—386	498	884	—0.32	0.41
R	0	1	0	1	2	—386	—102	488	—0.64	—0.17
J	0	1	1	1	2	214	—102	316	0.36	—0.17
J	0	1	1	2	3	14	98	112	0.03	0.25
R	0	1	2	3	5	94	18	112	0.39	0.08
J	0	1	2	4	7	— 43	— 16	59	—0.25	—0.09
R	0	1	4	7	12	14	— 2	16	0.14	—0.02
R	0	1	6	11	19	— 7	— 7	14	—0.12	—0.11
R	0	1	10	18	31	0.8	— 5.2	6.0	0.02	—0.13
J	0	1	11	20	34	1.9	3.9	5.8	0.05	0.11
J	0	1	17	31	53	1.4	— 0.1	1.5	—0.06	—0.003

(d) **Other Scales**

A_i	B_i	C_i	**Error** A_i	**Error** B_i	**Total Error**
5	10	17	—33	4	37
7	13	22	— 4	7	12

In a general mixed expansion, the choice of divisor at each stage can be arbitrary. The formulae for U_{i+1}, V_{i+1}, W_{i+1} are the same as for the Jacobi or the reversed expansion depending on whether one divides by U_i or V_i at step i. However, new formulae for A_i, B_i, C_i are needed.

Let J denote a Jacobi step and R a reversed step. Let $k = k(i)$ denote the number of steps between the present step, i, and the last previous J step. Let $k=i$ if there have been no J steps. Let S_i stand for either A_i, B_i, or C_i. Then define S_i by the following recursion formula,

$$S_i = p_i\, S_{i-k-3} + S_{i-2} + q_i\, S_{i-1} \quad \text{if step } i \text{ is } R \tag{5.1}$$

$$S_i = S_{i-k-3} + p_i\, S_{i-2} + q_i\, S_{i-1} \quad \text{if step } i \text{ is } J\ , \tag{5.2}$$

with the same initial conditions as for the Jacobi expansion. It is still clear that $C_i \to \infty$ as $i \to \infty$. The following identities are useful and can be proved by induction.

$$\begin{bmatrix} U_0 \\ V_0 \\ W_0 \end{bmatrix} = \begin{bmatrix} A_{i-k-3} & A_{i-2} & A_{i-1} \\ B_{i-k-3} & B_{i-2} & B_{i-1} \\ C_{i-k-3} & C_{i-2} & C_{i-1} \end{bmatrix} \begin{bmatrix} U_i \\ V_i \\ W_i \end{bmatrix} \tag{5.3}$$

If $i=0$, then $k=0$ and the formulae follow from the initial conditions. Suppose they hold at step i. If step i is J then from (3.3) and (5.2)

$$\begin{aligned} U_i S_{i-k-3} + V_i S_{i-2} + W_i S_{i-1} &= \\ &= \begin{bmatrix} U_{i+1} \\ V_{i+1} \\ W_{i+1} \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & p_i & q_i \end{bmatrix} \begin{bmatrix} S_{i-k-3} \\ S_{i-2} \\ S_{i-1} \end{bmatrix} \\ &= U_{i+1} S_{i-2} + V_{i+1} S_{i-1} + W_{i+1} S_i \end{aligned}$$

which is correct because at the $(i+1)$-th step k becomes 0.

Similarly, if step i is R then from (4.3) and (5.1),

$$\begin{aligned} & U_i S_{i-k-3} + V_i S_{i-2} + W_i S_{i-1} \\ &= (U_{i+1} + p_i W_{i+1}) S_{i-k-3} + W_{i+1} S_{i-2} + (V_{i+1} + q_i W_{i+1}) S_{i-1} \\ &= U_{i+1}\, S_{(i-2)-(k+1)} + V_{i+1}\, S_{i-1} + W_{i+1}\, S_i \end{aligned}$$

which is correct because at the $(i+1)$-th step k becomes $k+1$. Hence (5.3).

BARBOUR [2] does not discuss the convergence properties of the reversed and slow mixed expansions, but the argument of theorem 1 can be extended to prove the following result.

THEOREM 2. *Let $U_0 < V_0 < W_0$ be positive numbers linearly independent over the rationals. Given an arbitrary sequence of Jacobi and reversed steps we can expand them in an infinite mixed expansion. Let A_i, B_i, C_i be defined as above. Then $\lim_{i\to\infty} A_i / C_i = U_0 / W_0$ and $\lim_{i\to\infty} B_i / C_i = V_0 / W_0$.*

PROOF. As in the proof for the Jacobi expansion we only need to show that $H_i = A_i - (U_0 / W_0) C_i$ and $K_i = B_i - (V_0 / W_0) C_i$ are bounded in absolute value. Letting $\sigma_{1,i} = V_i/U_i$, $\sigma_{2,i} = W_i/U_i$, $\tau_{1,i} = U_i/V_i$, $\tau_{2,i} = W_i/V_i$ then we still know that $\sigma_{2,i}>1$, $\sigma_{2,i}/\sigma_{1,i}>1$, $\tau_{2,i}>1$, $\tau_{2,i}/\tau_{1,i}>1$ since $U_i<W_i$ and $V_i<W_i$ for all $i\geqslant 0$.

There are two cases to consider.

1. Suppose step i is J. Then $H_i=H_{i-k-3}+p_iH_{i-2}+q_iH_{i-1}$, and $H_i=-1/\sigma_{2,i+1}[H_{i-2}+\sigma_{1,i+1}H_{i-1}]$ or $H_i=-1/\tau_{2,i+1}[\tau_{1,i+1}H_{i-2}+H_{i-1}]$ depending on whether step $i+1$ is J or R. If H_{i-2}, H_{i-1} have opposite signs, then

$$|H_i| \leqslant \max\{|H_{i-2}|, |H_{i-1}|\}.$$

If H_{i-2}, H_{i-1} have the same sign, then H_i has the opposite sign and $|H_i| \leqslant |H_{i-k-3}|$.

2. Suppose step i is R. Then $H_i= p_iH_{i-k-3}+H_{i-2}+q_iH_{i-1}$, and $H_i=-1/\sigma_{2,i+1}[H_{i-k-3}+\sigma_{1,i+1}H_{i-1}]$ or $H_i=-1/\tau_{2,i+1}[\tau_{1,i+1}H_{i-k-3}+H_{i-1}]$ depending on whether step $i+1$ is J or R. If H_{i-k-3}, H_{i-1} have opposite signs, then

$$|H_i| \leqslant \max\{|H_{i-k-3}|, |H_{i-1}|\}.$$

If H_{i-k-3}, H_{i-1} have the same sign, then H_i has the opposite sign and $|H_i| \leqslant |H_{i-2}|$. Therefore, in any case

$$|H_i| \leqslant \max\{|H_{i-k-3}|, |H_{i-2}|, |H_{i-1}|\}.$$

Thus the $|H_i|$ sequence is bounded. Similar reasoning shows that the $|K_i|$ sequence is bounded and so the theorem follows. □

The proof of the above theorem does not depend on the particular sequence of Jacobi and reversed steps used, so the convergence of the Jacobi, reversed, and slow mixed expansions follow as special cases.

Note that we could reverse the order of U_0 and V_0 in Section 3 without affecting the validity of the Jacobi expansion. The effect would be the same as using a mixed expansion with one R step followed thereafter by J steps under the original order.

The slow mixed expansion of this section has been devised to slow the growth of the denominator C_i. Alternatively we could divide by the smaller of U_i and V_i at each step in order to speed the growth of C_i. We shall not explore this possibility further here.

6. Discussion of musical scales

Tables 1(a), 1(b) and 1(c) give the results of the Jacobi, reversed, and slow mixed expansions applied to the numbers $U_0=\log_2 5/4 \approx 0.3219$, $V_0=\log_2 3/2 \approx 0.5850$, $W_0=1$. Here C_i represents the number of notes in an octave, A_i the number of notes in the major third, and B_i the number of notes in the fifth. In order for A_i and B_i to be the best approximations for the denominator C_i we must have

$$|H_i| = |A_i-(U_0/W_0)C_i| < \tfrac{1}{2} \text{ and } |K_i| = |B_i-(V_0/W_0)C_i| < \tfrac{1}{2}.$$

For all of the interesting cases these inequalities are easily satisfied.

The errors in A_i/C_i and B_i/C_i are measured in cents,

$$\text{Error } A_i = 1200(A_i/C_i - U_0/W_0)$$

$$\text{Error } B_i = 1200(B_i/C_i - V_0/W_0) .$$

The total error is taken to be $|\text{ Error } A_i \,| + |\text{ Error } B_i \,|$.

The scales from these expansions include many of the important scales proposed by musical theorists and several scales in use by various non-western cultures. The following comments are taken from Barbour ([2] and [3], Chapter 6) who discusses these and other scales in more detail.

Two of the scales with fewer than 12 notes are worth mentioning. According to Barbour, Javanese music is based on an evenly-tempered 5-note scale, and Siamese music, on an evenly-tempered 7-note scale.

In western 12-note music there are 5 whole steps and 2 (diatonic) half-steps in the octave. If each whole step is split into a diatonic and a chromatic half-step, there are 7 diatonic and 5 chromatic half-steps. In an evenly-tempered tuning the ratio between these two kinds of half-step is taken as 1:1. If instead one takes the ratio to be 2:1 one gets $7\times2+5\times1=19$ notes in the octave. Thus, one can get a non-evenly-tempered 12-note scale by taking 12 notes out of an evenly-tempered 19 scale.

Other important scales in Table 1(c) which can be interpreted in this way are those with 31 notes (ratio 3:2) and 53 notes (ratio 4:5).

Arabian music is based on a 17-note evenly-tempered scale. This scale has a good fifth (within 4 cents), but the major third is very flat being about midway between a true major third and minor third. The poorness of the third probably explains why it does not appear in the expansions.

The 22-note scale is one important scale missing from these tables though it appears under a more general mixed expansion. It is interesting to note that both the 19 and 22-note scales form better approximations in terms of total error than the 25-note scale in the Jacobi expansion. Thus from this point of view, the convergents of the Jacobi expansion are not necessarily best possible approximations. In this respect Jacobi ternary continued fractions are weaker than ordinary continued fractions, because in ordinary continued fractions one gets a best possible approximation at every step.

From the musical point of view, the accuracy of the pure fifth is more important than the accuracy of the major third. The concept of total error does not take this feature into account. The 12-note scale has a better pure fifth than any evenly-tempered scale with fewer than 41 notes.

The only possible systems of multiple division of the octave which could have any practical significance are the 19 and 22-note scales. Any more notes than that would make an instrument extremely unwieldy to play. Further, as it does not seem likely that the 19- or 22-note scales will come into widespread acceptance, most music seems destined to remain in the evenly-tempered 12-note scale.

References

1. A. Baker (1984). *A Concise Introduction to the Theory of Numbers,* Cambridge University Press.
2. J.M. Barbour (1948). Music and ternary continued fractions. *Am. Math. Monthly 55,* 545-555.
3. J.M. Barbour (1951). *Tuning and Temperament,* Michigan State College Press, East Lansing.
4. P.H. Daus (1922). Normal ternary continued fraction expansions for the cube roots of integers. *Am. Jour. of Math. 44,* 279-296.
5. G.H. Hardy, E.M. Wright (1960). *An Introduction to the Theory of Numbers,* Oxford University Press.
6. C.G.J. Jacobi (1891). Allgemeine Theorie der Kettenbruchähnlichen Algorithmen, in welchen jede Zahl aus drei Vorhergehenden gebildet wird. *Gesammelte Werke, Band 6, Reimer,* Berlin, 385-426.
7. V.O. Perron (1907). Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus. *Math. Annalen 64,* 1-76.
8. V.O. Perron (1935). Ein Satz über Jacobi-ketten zweiter Ordnung. *Annali di Pisa 4,* 133-138.
9. K. Vaisälä (1929). Zur Theorie des Jacobischen Kettenbruchalgorithmus Zweiter Ordnung, *Lindelöf-Festschrift, Helsinki.*

Multiple Correspondence Analysis

and

Ordered Latent Structure Models

B.F. Schriever

Nederlandse Philips Bedrijven B.V., Centre for Quantitative Methods
P.O. Box 218, 5600 MD Eindhoven

This paper discusses application of multiple correspondence analysis in ordered latent structure models. Such models are frequently used in psychological measurement theory to analyse ability (or attitude) tests, e.g. intelligence tests. The models considered are related to those of Mokken. It turns out that, under realistic assumptions, multiple correspondence analysis orders the questions (items) in the test according to their difficulty and orders individuals according to their ability (or attitude).

1. Introduction

Multiple correspondence analysis (abbreviated MCA) is a modern statistical technique for describing the association between categorical (i.e. discrete) variables. The technique is commonly used to analyse large data sets. It gives insight into the complex dependence structure of such data sets by making plots. MCA has proved to be an important and useful tool for analysing the association which is present in data sets with many variables. In this paper we discuss MCA in a more unusual situation and we need not consider its graphical representation. We apply MCA in the analysis of ordered latent structure models. Such models are developed for the following situation which frequently arises in e.g. psychology and medicine. In a population individuals must be ordered according to their value on an *unobservable* characteristic (e.g. intelligence, knowledge of a subject, attitude in a given context, a specific disease). For this purpose responses on a set of variables related to the characteristic are collected for each individual (e.g. an intelligence test). We restrict attention to the simple case in which there is only one such characteristic, called the *latent variable,* and in which the collected response variables, called *items,* are dichotomous (i.e. have only two response categories; for example 'correct' and 'wrong'). The set of all items is called the *test.* Since the characteristic of interest is often hard to separate from other characteristics, the assumption that responses on the items (i.e. the response variables) systematically depend on only one latent variable is for most applications more restrictive than the dichotomy assumption.

The paper is organized as follows. Section 2 reviews the definition of the technique MCA. In Section 3 we introduce latent structure models with ordered items. These ordered models are special cases of the models introduced by MOKKEN [17]. Our main result is given in Section 4. It demonstrates that the ordering of items is reflected in the MCA scores. This implies that MCA orders the individuals according to their latent value and orders the items according to their difficulty. In the last section it is shown that most well-known examples of latent structure models possess the orderings of Section 3. GIFI ([5], Chap. 9) already noted the ordering property of MCA for these specific examples, but proofs were not given.

The result presented in this paper also appeared in SCHRIEVER [23].

2. MULTIPLE CORRESPONDENCE ANALYSIS

Let $X_1, X_2, ..., X_k$ be categorical random variables. The technique MCA seeks k real valued functions $\phi_{11}, \phi_{21}, ..., \phi_{k1}$, defined on the categories (possible values) of $X_1, X_2, ..., X_k$ respectively, such that the first principal component of the correlation matrix of $\phi_{11}(X_1), \phi_{21}(X_2), ..., \phi_{k1}(X_k)$ has maximal variance. This principal component is called the first MCA component. It describes the most informative part of the variation between the categorical variables. Clearly, it is no restriction to assume that the derived variables $\phi_{l1}(X_l)$ have expectation zero and variance unity, for $l = 1, ..., k$. Subsequently, MCA seeks a second component which has maximal variance but which is uncorrelated with the first. This procedure is continued with a third component, a fourth component, etc. until no new component which is uncorrelated with the previous components can be found.

DEFINITION. *The $t-th$ MCA component is the linear combination of transformed variables*

$$Y_t = \sum_{l=1}^{k} \alpha_{lt} \phi_{lt}(X_l)$$

for which $\mu_t = Var(Y_t)$ is maximal subject to

$$E\phi_{lt}(X_l) = 0, \quad Var(\phi_{lt}(X_l)) = 1 \textit{ for } l = 1, ..., k,$$

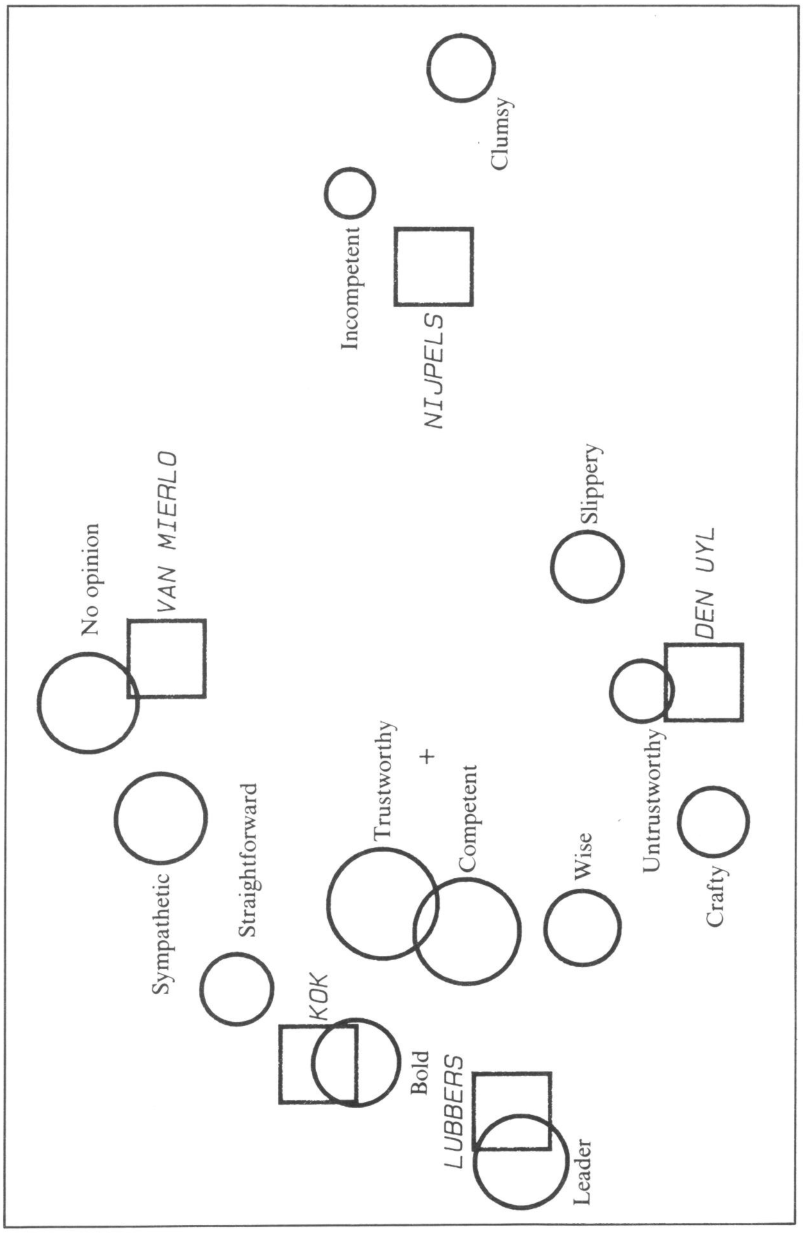
and the normalization constraint

$$\sum_{l=1}^{k} \alpha_{lt}^2 = 1, \tag{2.1}$$

$$Corr(Y_t, Y_s) = 0 \textit{ for } s = 1, ..., t-1.$$

The MCA *solution* consists of all $k+1$ tuples $(\mu_t, \alpha_{1t}\phi_{1t}(X_1), ..., \alpha_{kt}\phi_{kt}(X_k))$ for $t = 1,2,...$. The value $\alpha_{lt}\phi_{lt}(x)$ is called the *category score* on the t-th MCA component of the category x of X_l; $l = 1, ..., k$; $t = 1,2,...$.

The present definition of MCA may depart from other definitions given in the literature with respect to the normalization of the variable weights



The illustration shows the result of a classical correspondence analysis. Respondents were asked to characterize each of a number of Dutch politicians. The picture shows the embedding of politicians and characteristics in a two-dimensional space which preserves as closely as possible chi-square type measures of distance derived from the data.

$\alpha_{1t},...,\alpha_{kt}$. Also, different names for this technique are used in the literature, e.g. homogeneity analysis (GIFI [5]) and first order correspondence analysis (HILL [9]).

It follows directly from the definition that

$$Var(Y_t) = \sum_{j=1}^{k} \sum_{l=1}^{k} \alpha_{jt}\alpha_{lt} Corr(\phi_{jt}(X_j), \phi_{lt}(X_l))$$

which means that MCA only considers the bivariate marginals of the k-dimensional probability distribution of $X_1,...,X_k$. It is well-known (cf. GIFI [5]; GREENACRE [6]; HILL [9]; LEBART et al. [13]; SCHRIEVER [23]) that an MCA solution always exists and can be obtained by solving a generalized eigenvalue problem of the super matrix containing all bivariate marginal probability distributions.

MCA can be seen as a generalization of principal component analysis to nominal variables. Moreover, when $X_1, X_2,..., X_k$ are all dichotomous, e.g. $0-1$ variables, then by the normalization (2.1), $\phi_{lt}(1) = ((1-\pi_l)/\pi_l)^{\frac{1}{2}}$ and $\phi_{lt}(0) = -(\pi_l/(1-\pi_l))^{\frac{1}{2}}$, where $\pi_l = P\{X_l = 1\} = 1-P\{X_l = 0\}$ for $l = 1,2,...,k$ and $t = 1,2,....$ (Note that the signs of $\phi_{lt}(1)$ and $\phi_{lt}(0)$ may be taken arbitrary but opposite.) Hence the variance of Y_t is only maximized with respect to the variable weights $\alpha_{1t}, \alpha_{2t},..., \alpha_{kt}$ for $t = 1,2,....$ Therefore, MCA in the dichotomous case is equivalent to finding the principal components of the covariance matrix of $\phi_{11}(X_1), \phi_{21}(X_2),..., \phi_{k1}(X_k)$, that is, of the correlation matrix of $X_1, X_2,..., X_k$.

For further properties, for different approaches and for applications of MCA consult DE LEEUW [2], GIFI [5], GREENACRE [6], LEBART et al. [13] and NISHISATO [20].

3. ORDERED LATENT STRUCTURE MODELS

The latent structure model we consider supposes that the responses of the individuals on the k dichotomous items (variables) $X_1, X_2,..., X_k$ can be accounted for, to a substantial degree, by one latent variable Z. It is assumed that conditionally on Z the items $X_1, X_2,..., X_k$ are stochastically independent. This assumption of *local independence* means that each individual responds independently on the items. This implies that *the (global) dependence structure between the items is caused and hence can completely be explained by variation in the latent variable.* Local independence is essential in latent structure models; if it does not hold then the latent variable cannot be distinguished from other interactions between the items.

Let the probability distribution function of Z be denoted by H; i.e. $H(z) = P\{Z \leq z\}$. Our results are not based on any assumption on H and thus hold for any (sampled) population. Let the response categories of each item be labeled with 1 ('correct') and 0 ('wrong'). The probability of a correct response on item X_l for an individual with latent value z is denoted by

$$\pi_l(z) = P\{X_l = 1 | Z = 1\} \text{ for } z \in \mathbb{R};\ l = 1,...,l.$$

It can be interpreted as the (local) difficulty of item X_l for this individual. The function $\pi_l(\cdot)$ is called the *trace line* of item X_l; $l = 1,...,k$. The unconditional probability of a correct response on item X_l,

$$\pi_l = P\{X_l = 1\} = \int_{\mathbf{R}} \pi_l(z)dH(z),$$

is the (global) difficulty of this item for the population. By local independence, the joint probability of correct responses to both item X_l and item X_j, $j \neq l$, for an individual with latent value z is given by $\pi_j(z)\pi_l(z)$. The unconditional joint probability of correct responses to both items is denoted by

$$\pi_{lj} = \int_{\mathbf{R}} \pi_l(z)\pi_j(z)dH(z) \quad \text{for } l,j = 1,...,k\,;\ l \neq j,$$

but we define $\pi_{ll} = \pi_l$ for $l = 1,...,k$. It is easily shown that the correlation between the items X_l and X_j equals

$$\sigma_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_j(1-\pi_j)\pi_l(1-\pi_l))^{\frac{1}{2}} \quad \text{for } l,j = 1,...,k.$$

The correlation matrix of the items $X_1, X_2,...,X_k$ is denoted by $\Sigma = (\sigma_{lj})$.

MOKKEN [17] imposes two natural conditions on the trace lines of the items in the test. First, he assumes for each item that the probability of a correct response increases as the individual has a higher latent value, i.e., for $l = 1,...,k$:

$$z_1 < z_2 \Rightarrow \pi_l(z_1) \leqslant \pi_l(z_2) \text{ and not almost everywhere } (dH)\text{--equality}; \qquad (3.1)$$

Secondly, Mokken assumes that if for one individual an item is more difficult than another item, then it must be more difficult for all individuals. In other words, the trace lines of the items may not cross each other. This means that the items in the test can be indexed such that

$$1 \leqslant l < j \leqslant k \Rightarrow \pi_l(z) \geqslant \pi_j(z) \text{ for all } z \text{ and not } dH-\text{ a.e. equality.} \qquad (3.2)$$

The items in the test are then indexed from easy to difficult. Tests satisfying (3.1) and (3.2) are called *doubly monotone*. More about interpretation and examples of doubly monotone tests can be found in MOKKEN [17]. In many examples, see for instance those of Section 5, double monotonicity typically occurs in combination with the following stronger ordering of trace lines,

$$z_1 < z_2,\ 1 \leqslant l < j \leqslant k \Rightarrow \pi_l(z_1)\pi_j(z_2) \geqslant \pi_l(z_2)\pi_j(z_1). \qquad (3.3)$$

In this case the trace lines in the test are said to be *totally positive of order* 2 (TP_2). Also, a similar TP_2 property with respect to the wrong responses,

$$z_1<z_2,\ 1\leqslant l<j\leqslant k\Rightarrow(1-\pi_l(z_1))(1-\pi_j(z_2))\geqslant(1-\pi_l(z_2))(1-\pi_j(z_1)) \qquad (3.4)$$

frequently holds.

The increasing property (3.1) implies that all items are positively correlated, because

$$\pi_{lj} - \pi_l\pi_j = \frac{1}{2}\int\int_{\mathbf{R}^2}(\pi_l(z_2)-\pi_l(z_1))(\pi_j(z_2)-\pi_j(z_1))dH(z_2)dH(z_1) > 0 \qquad (3.5)$$

for $l,j = 1,...,k$. Thus the correlations σ_{lj} for $l,j = 1,...,k$ are even strictly positive. Moreover, it trivially follows from (3.2) that

$$1 \leqslant l < j \leqslant k \Rightarrow \pi_l > \pi_j \text{ and } \pi_{li} \geqslant \pi_{ji} \text{ for } i \neq j. \tag{3.6}$$

Large departures from double monotonicity violate (3.5) and (3.6) and might be detected by inspection of these properties. Notice that (3.5) and (3.6) only concern properties of the observable items.

All these assumptions concern the underlying (population) model. We make some remarks on the statistical problems which arise when one has only a finite sample of real data at the end of the next section.

4. Analysis of the model with MCA

Analysis with MCA of the latent structure model described in the previous section may be motivated by the interpretation of this technique and by the main result of this section. Recall that in the dichotomous case the first MCA component Y_1 equals the first principal component of the correlation matrix Σ of the items $X_1, X_2, ..., X_k$. Therefore, Y_1 'best explains' the dependence structure between the items among all linear combinations of items. Since the latent variable completely explains this dependence structure, Y_1 can be interpreted as the linear combination of items which best fits the latent variable Z in this sense. So the model will be analysed using the correct and wrong category scores, $\gamma_{l1} = \alpha_{l1}\phi_{l1}(1)$ and $\omega_{l1} = \alpha_{l1}\phi_{l1}(0)$ for $l = 1,...,k$, on the first MCA component. It follows from Section 2 that

$$\gamma_{l1} = ((1-\pi_l)/\pi_l)^{\frac{1}{2}}\,\alpha_{l1} \text{ and } \omega_{l1} = -(\pi_l/(1-\pi_l))^{\frac{1}{2}}\,\alpha_{l1}$$

for $l = 1,...,k$, where $\alpha_1 = (\alpha_{11},...,\alpha_{k1})^T$ is the eigenvector of Σ corresponding to the largest eigenvalue. (The superscript T denotes the transposition of a vector.)

Now suppose that a subset of items in the test satisfies the double monotonicity and TP_2 conditions. Then these items possess strong orderings with respect to their (local and global) difficulties. The next theorem shows that these orderings are reflected in the correct and wrong category scores of these items, even when the remaining items do not match the orderings of the items in the subset.

Theorem. *Suppose the test consists of m items which all satisfy (3.1) with k replaced by m. Furthermore, suppose k of the items, which without loss of generality can be taken as the first k, can be indexed such that (3.2) and (3.3) hold. Then the correct scores of these k items satisfy*

$$0 < \gamma_{11} < \gamma_{21} <...< \gamma_{k1}. \tag{4.1}$$

Similarly, if (3.1), (3.2) and (3.4) hold for these first k items, then

$$\omega_{11} < \omega_{21} <...< \omega_{k1} < 0. \tag{4.2}$$

PROOF. Let S denote the $k \times k$ lower triangular matrix with unit elements on and below the diagonal and all other elements zero. Its inverse S^{-1} is the matrix with unit elements on the main diagonal, with elements -1 adjacent and below the diagonal (i.e. on the first sub diagonal) and all other elements zero. Denote by T the $m \times m$ block diagonal matrix with diagonal blocks S and the $(m-k)\times(m-k)$ identity matrix I. Then its inverse T^{-1} is a block diagonal matrix with diagonal blocks S^{-1} and I.

Note that the vector $\gamma_1 = (\gamma_{11},...,\gamma_{m1})^T$ is an eigenvector corresponding to the largest eigenvalue of the matrix C with elements $c_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_l(1-\pi_j))$ for $l,j = 1,...,m$. Since T is non-singular, γ_1 is an eigenvector corresponding to the largest eigenvalue of C iff $d = T^{-1}\gamma_1$ is an eigenvector corresponding to the largest eigenvalue of $D = T^{-1}CT$.

Under the conditions of the theorem, all elements of D turn out to be positive (i.e. larger than or equal to zero) and all elements of D^2 even turn out to be strictly positive. This can be verified as follows. The elements of the matrix $T^{-1}C = B = (b_{lj})$ equal

$$b_{lj} = (\pi_{lj} - \pi_l\pi_j)/(\pi_l(1-\pi_j)) \quad \text{for} \quad l = 1,\ k+1,...,m\,;\ j = 1,...,m,$$

$$b_{lj} = (\pi_{l-1}\pi_{lj} - \pi_l\pi_{l-1j})/(\pi_l\pi_{l-1}(1-\pi_j)) \quad \text{for} \quad l = 2,...,k\,;\ j = 1,...,m.$$

Since (3.1) holds for all m items, it follows that all correlations are strictly positive and hence $b_{lj} > 0$ for $l = 1,k+1,...,m$; $j = 1,...,m$. Furthermore, by (3.4), $\pi_l^{-1}\pi_l(z)$ can be interpreted as a density (with respect to the measure dH) which has the monotone likelihood ratio or TP_2 property. Therefore, since $\pi_j(z)$ is increasing in z for each $j = 1,...,m$, it follows from LEHMANN [14], p. 74) or KARLIN ([10], p. 22) that $\int \pi_l^{-1}\pi_l(z)\pi_j(z)dH(z)$ is increasing in l. Thus, $\pi_{l-1}\pi_{lj} - \pi_l\pi_{l-1j} \geqslant 0$ and hence $b_{lj} \geqslant 0$ for $l = 2,...,k$ and $j \neq l-1,l$. Obviously, $b_{ll} > 0$ and $b_{ll-1} < 0$ for $l = 2,...,k$. So the matrix B has positive elements except for b_{ll-1}; $l = 2,...,k$. But by (3.2), $\pi_{l-1} > \pi_l$ which implies that $b_{ll-1} + b_{ll} > 0$ for $l = 2,...,k$. Therefore, $D = BT$ has positive elements. Moreover, since $b_{lj} > 0$ for $l = 1,k+1,...,m$ and $j = 1,...,m$ and since $b_{ll-1} + b_{ll} > 0$ for $l = 2,...,k$, it follows that the elements in the first row of D and the elements in the first column of D are strictly positive. This implies that all elements of D^2 are strictly positive.

Application of the Perron-Frobenius theorem (cf. GANTMACHER [4], p. 53 or RAO [21], p. 46) yields that the eigenvector $d = (d_1,...,d_m)^T$ corresponding to the largest eigenvalue of D^2 (or of D) has strictly positive components. Since $d = T^{-1}\gamma_1$ or equivalently $d_l = \gamma_{l1}$ for $l = 1,k+1,...,m$ and $d_l = \gamma_{l1} - \gamma_{l-11}$ for $l = 2,...,k$, the result (4.1) follows. The proof of (4.2) is similar. □

The conditions (3.2), (3.3) and (3.4) of the theorem can be relaxed; see SCHRIEVER [23].

This result shows that the MCA correct and wrong category scores reflect the difficulties of the items. Since these scores do not depend on the order in which items are presented to MCA, this ordering property can be used, in combination with (3.5) and (3.6), for a first investigation of the model assumptions. Moreover, the theorem suggests that ordering the individuals according to their MCA test score Y_1 is reasonable: responding to a difficult item correctly yields a large contribution to this test score and responding to it wrongly does not cost much, whereas for an easy item it is the other way around. The test score Y_1 is a weighted sum of items with small weights for items which are less related to the latent variable than other items.

Analysis with MCA is an alternative to the method proposed by MOKKEN [17] in which individuals are ordered according to the unweighted sum of correct responses. It is unknown, however, in which cases which method actually works better. An advantage of the MCA approach is that our results can be generalized, in a natural fashion, to the case where items have three or more response categories. Such a generalization of Mokken's method, as discussed by MOLENAAR [19], is quite complicated and less natural.

In practice the probabilities π_l and π_{jl} for $j,l = 1,...,m$ have to be estimated by the relative frequencies of correct responses. Although the MCA scores based on these estimates need not reflect the difficulties of the items even when the underlying model satisfies the assumptions of the theorem, one would expect the total score Y_1 to reflect the appropriate ordering of the individuals quite well. It is, however, difficult to derive precise and useful statistical properties of such qualitative aspects.

5. EXAMPLES OF MODELS

Latent structure models for dichotomous variables studied in the literature (e.g. ANDERSEN [1]; FISCHER [3]; LORD and NOVICK [16] are commonly of parametric form, that is, the functional form of the trace lines is specified. Often there is, however, no evidence that the specificd functional form is actually present in the test at hand. The parametric examples below generally satisfy the double monotonicity and TP_2 conditions and, therefore, analysis both with MCA and with Mokken's method is legitimate. These examples are also discussed in MOKKEN [17] and MOLENAAR [18]; the ordering property of MCA under these models are mentioned, but not proved in GIFI [5].

In *Guttman's model* the responses on the items are deterministic functions of the latent variable. The trace lines are given by

$$\pi_l(z) = \begin{cases} 1 & \text{if } z \geqslant \delta_l \\ 0 & \text{if } z < \delta_l \end{cases} \quad \text{for } l = 1,...,k,$$

where the item parameters satisfy $\delta_1 < \delta_2 < ... < \delta_k$. In this model an individual cannot respond correctly to a difficult item and wrongly to an easier item. Hence a perfect analysis is possible. Double monotonicity and TP_2 of trace lines are easily verified and thus by the theorem the correct and wrong category scores of the items (on the first MCA component) increase as the item becomes more difficult. Moreover, it is demonstrated in SCHRIEVER [23] that the correct and wrong category scores on higher MCA component are oscillating functions of the item difficulty. The practical relevance of these stronger ordering properties is, however, limited. (Slightly weaker oscillatory properties for the principal components $\alpha_1,...,\alpha_k$ of Σ are proved in GUTTMAN [7] and interpreted in GUTTMAN [8].)

A somewhat more realistic generalization of the previous model is the *latent distance model* of LAZARSFELD and HENRY [12]. The trace lines of this model satisfy

$$\pi_l(z) = \begin{cases} 1-\zeta_l & \text{if } z \geqslant \delta_l \\ \epsilon_l & \text{if } z < \delta_l \end{cases} \quad \text{for } l = 1,...,k,$$

where $\delta_1 < \delta_2 <...< \delta_k$ and $\epsilon_l < 1-\zeta_l$ for $l = 1,...,k$. If $\epsilon_l > 0$ and $\zeta_l > 0$ for $l = 1,...,k$, then double monotonicity and TP_2 cannot hold simultaneously. But the weaker conditions for our main theorem are satisfied when $\pi_{l-1} > \pi_l$, $\pi_l \epsilon_{l-1} \geqslant \pi_{l-1}\epsilon_l$ and $\pi_l(1-\zeta_{l-1}) \geqslant \pi_{l-1}(1-\zeta_l)$ for $l = 2,...,k$ and hence (4.1) and (4.2) remain valid (cf. SCHRIEVER [23]).

In the *linear model* of Lazarsfeld the trace lines satisfy $\pi_l(z) = a_l z + b_l$ provided $0 \leqslant a_l z + b_l \leqslant 1$ for $l = 1,...,k$. The conditions (3.1) and (3.2) are for instance satisfied when $a_{l-1} \leqslant a_l$ and $b_{l-1} \geqslant b_l$ for $l = 2,...,k$. The TP_2 conditions (3.3) and (3.4) hold when $a_{l-1} b_l \leqslant a_l b_{l-1}$ and $a_{l-1}(1+b_l) \geqslant a_l(1+b_{l-1})$ for $l = 2,...,k$.

RASCH [22] developed a model in which the unweighted sum of all correct responses is sufficient for Z. The trace lines are given by $\pi_l(z) = z/(z+\delta_l)$ or $= 0$ as $z \geqslant 0$ or $z < 0$, where $0 < \delta_1 <...,\delta_k$. This model is a special case of a model considered by Birnbaum (cf. Chap. 17-20, LORD and NOVICK [16]).). In Rasch's model the unweighted sum of correct responses 'uniformly best discriminates' the individuals (cf. MOKKEN [17], p. 141). Double monotonicity and TP_2 of trace lines is easily verified.

The last example consists of models based on shifts in distribution functions. For an univariate distribution function F the trace lines are defined by $\pi_l(z) = F(z-\delta_l)$ for $l = 1,...,k$ and $\delta_1 < \delta_2 <...< \delta_k$. Double monotonicity is obvious. The TP_2 conditions hold when the density p of F (with respect to some measure) is log concave. Special choices of such distributions F yield well-known models, e.g. degenerate distribution (Guttman's model), logistic distribution (Rasch's model), normal distribution (models of LAWLEY [11] and LORD [15]). Other examples of such distribution functions are the gamma, Poisson and binomial distribution function.

REFERENCES

1. E.B. ANDERSEN (1980). *Discrete Statistical Models with Social Science Applications,* North-Holland Publ. Co; Amsterdam.
2. J. DE LEEUW (1984). Statistical properties of multiple correspondence analysis. Paper presented at the conference on New Multivariate Methods in Statistics. *The 1984 Joint Summer Research Conference Series in Mathematical Sciences,* June 10-16, Bowdoin College, Brunswick, Maine. (Report RR-84-06, Dept. of Data Theory, University of Leiden).
3. G.H. FISCHER (1974). *Einführung in die Theorie der Psychologischer Tests,* Huber, Bern.

4. F.R. GANTMACHER (1977). *Matrix Theory, Vol II,* Chelsea, New York.
5. A. GIFI (1981). *Non-Linear Multivariate Analysis,* Dept. of Data Theory, University of Leiden.
6. M.J. GREENACRE (1984). *Theory and Applications of Correspondence Analysis,* Academic Press, New York.
7. L. GUTTMAN (1950). The principal components of scale analysis. S.A. STOUFFER et al. *Measurement and Prediction,* Princeton University Press, Princeton, pp. 312-361.
8. L. GUTTMAN (1954). The principal components of scalable attitudes. P.F. LAZARSFELD (ed.). *Mathematical Thinking in the Social Sciences,* Free Press, Glencoe, pp. 216-257.
9. M.O. HILL (1974). Correspondence analysis: a neglected multivariate method. *Appl. Statist. 23,* 340-354.
10. S. KARLIN (1968). *Total Positivity I,* Stanford University Press, Stanford.
11. D.N. LAWLEY (1943). On problems connected with item selection and test construction. *Proc. Royal Society of Edinburgh 61,* 273-287.
12. P.F. LAZARSFELD, N.W. HENRY (1968). *Latent Structure Analysis,* Houghton-Mifflin, Boston.
13. L. LEBART, A. MORINEAU, N. TABARD (1977). *Techniques de la Description Statistique,* Dunod, Paris. (English translation: L. LEBART, A. MORINEAU, K.M. WARWICK (1984). *Multivariate Descriptive Statistical Analysis,* Wiley, New York.)
14. E.L. LEHMANN (1959). *Testing Statistical Hypotheses,* Wiley, New York.
15. F.M. LORD (1952). *A Theory of Test Scores, Psychometric Monograph No. 7.*
16. F.M. LORD, M.R. NOVICK (1968). *Statistical Theories of Mental Test Scores,* Addison Wesley, Reading (Mass).
17. R.J. MOKKEN (1971). *A Theory and Procedure of Scale Analysis,* Mouton, The Hague.
18. I.W. MOLENAAR (1982). Mensen die het beter meten. *Kwantitatieve Methoden 5,* 3-29. (Dutch Statistical Society)
19. I.W. MOLENAAR (1982). Mokken scaling revisited. *Kwantitatieve Methoden 8,* 145-164. (Dutch Statistical Society)
20. S. NISHISATO (1980). *Analysis of Categorical Data: Dual Scaling and its Applications,* University of Toronto Press, Toronto.
21. C.R. RAO (1973). *Linear Statistical Inference and its Applications,* Wiley, New York.
22. G. RASCH (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Danish Institue for Educational Research, (Reprinted 1980: University of Chicago Press.)
23. B.F. SCHRIEVER (1986). *Order Dependence, CWI Tract 20.* Centre for Mathematics and Computer Science, Amsterdam.

Book Review

Sara van de Geer
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

D. POLLARD (1984). *Convergence of Stochastic Processes,* Springer Series in Statistics, Springer Verlag, New York etc.

One of the reasons for studying convergence of probability measures on metric spaces is to obtain useful approximation theorems in statistics. For instance, suppose $X_1, X_2, \ldots$ are independent observations from a probability distribution P on $\mathbb{R}$. Let P_n be the empirical measure, i.e. P_n puts equal mass n^{-1} at each of the first n sample points $X_1, \ldots, X_n$. For example, $P_n((-\infty,a]) = n^{-1} \cdot \{\# X_i \leqslant a,\ i \leqslant n\}$, and for general sets A, $P_n(A) = n^{-1} \cdot \{\# X_i \in A,\ i \leqslant n\}$. By the Glivenko-Cantelli Theorem

$$\lim_{n \to \infty} \sup_{a \in \mathbb{R}} |P_n((-\infty,a]) - P((-\infty,a])| = 0, \tag{1}$$

almost surely. The uniformity in a implies the almost sure convergence of certain functionals of $P_n((-\infty,\cdot])$. A more refined result says that the empirical process

$$\nu_n((-\infty,a]) := \sqrt{n}(P_n((-\infty,a]) - P((-\infty,a]))$$

converges weakly to a Brownian bridge. This suggests e.g. an approximation for the level of the Kolmogorov-Smirnov test.

A famous book on weak convergence of $\nu_n(\cdot)$ as random element of a space D of functions on $\mathbb{R}$ is Billingsley's *Convergence of Probability Measures* [1]. Pollard organizes some of the more recent theory on convergence of random elements of $\mathfrak{X}$, $\mathfrak{X}$ being some space of functions with domain $\mathfrak{F}$, where $\mathfrak{F}$ is a class of functions on a probability space $(S, \mathbb{S}, P)$. The emphasis is on empirical processes indexed by functions $f \in \mathfrak{F}$ — or sets $A \in \mathfrak{A}$ —, which are indeed elements of such an abstract space $\mathfrak{X}$.

The book covers a broad field, from which everyone can choose his or her favourite subject. I shall mainly highlight the extension of the Glivenko-Cantelli Theorem to vector-valued random variables (Chapter II of the book). Here, the concept of empirical processes indexed by sets already comes up naturally. Let $X_1, X_2, \ldots$ be independently sampled from a distribution P on $\mathbb{R}^d$. Let $\mathcal{Q}$ be a class of measurable subsets $A \subset \mathbb{R}^d$. By the law of large numbers, for each measurable A

$$|P_n(A) - P(A)| \to 0 \text{ almost surely,}$$

i.e. relative frequencies converge to probabilities. The problem is now to find conditions on the class $\mathcal{Q}$ such that the almost sure convergence holds uniformly in $A \in \mathcal{Q}$.

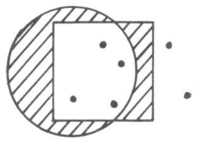
EXAMPLE 1. In the Glivenko-Cantelli Theorem, $d=1$, $\mathcal{Q} = \{(-\infty, a],\ a \in \mathbb{R}\}$ and (1) can be written as

$$\sup_{A \in \mathcal{Q}} |P_n(A) - P(A)| \to 0 \text{ almost surely.}$$

Obviously, if $\mathcal{Q}$ is finite, then the law of large numbers holds uniformly over $\mathcal{Q}$. Now, no matter how large $\mathcal{Q}$, the number of different $A \in \mathcal{Q}$ one can distinguish from a sample of size n is always at most 2^n. Formally, if we define

$$\Delta^{\mathcal{Q}}(X_1, \ldots, X_n) := \# \{A \cap \{X_1, \ldots, X_n\} : A \in \mathcal{Q}\},$$

then $\Delta^{\mathcal{Q}}(X_1, \ldots, X_n) \leqslant 2^n$. We say that two sets A and $\tilde{A}$ differ for P_n if $P_n(A \Delta \tilde{A}) \neq 0$, where $A \Delta \tilde{A}$ is the symmetric difference $(A \cap \tilde{A}^c) \cup (A^c \cap \tilde{A})$



$$P_n(A \Delta \tilde{A}) = 0$$

Let $\{A_1, \ldots, A_m\}$ be such that for each $A \in \mathcal{Q}$, $P_n(A \Delta A_j) = 0$ for some $A_j \in \{A_1, \ldots, A_m\}$. Then $\Delta^{\mathcal{Q}}(X_1, \ldots, X_n)$ is the smallest value of m for which such a collection $\{A_1, \ldots, A_m\}$ exists.

EXAMPLE 1 continued. Take

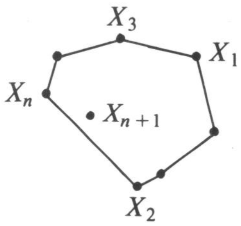
$$A_j = (-\infty, X_{(j)}),\ j = 1, 2, \ldots, n+1$$

where $X_{(1)} \leqslant X_{(2)} \leqslant \ldots \leqslant X_{(n)}$ are the order statistics, and $X_{(n+1)} = \infty$. Then for $A = (-\infty, a]$, $P_n(A \Delta A_j) = 0$ for $X_{(j-1)} \leqslant a < X_{(j)}$. Thus $\Delta^{\mathcal{Q}}(X_1, \ldots, X_n) \leqslant n+1$.

If $\Delta^{\mathcal{Q}}(X_1, \ldots, X_n) = 2^n$, one says that $\mathcal{Q}$ *shatters* $\{X_1, \ldots, X_n\}$.

Example 2. Let $\mathcal{C}$ be the collection of all closed convex sets in $\mathbb{R}^2$. If $X_1,...,X_n$ all lie on their convex hull, then $\Delta^{\mathcal{C}}(X_1,...,X_n) = 2^n$. And if X_{n+1} falls inside this convex hull, then any closed convex set containing $\{X_1,...,X_n\}$ must also contain X_{n+1}, i.e.

$$\{X_1,...,X_n\} \notin \{A \cap \{X_1,...,X_n,X_{n+1}\}: A \in \mathcal{C}\}.$$



So $\mathcal{C}$ shatters $\{X_1,...,X_n\}$, but does not shatter $\{X_1,...,X_n, X_{n+1}\}$.

The rate of growth of $\Delta^{\mathcal{C}}(X_1,...,X_n)$ determines whether or not the uniform law of large numbers holds.

Example 3. Let $\mathcal{C}$ be the collection of all finite subsets of [0,1], and P the uniform distribution. Then $P(A) = 0$ for all $A \in \mathcal{C}$, whereas $P_n(\{X_1,...,X_n\}) = 1$ for all n. Thus

$$\sup_{A \in \mathcal{C}} |P_n(A) - P(A)| = 1.$$

Moreover, $\Delta^{\mathcal{C}}(X_1,...,X_n) = 2^n$ for all n.

Vapnik and Chervonenkis [3] show that (neglecting some measurability problems),

$$n^{-1}\log\Delta^{\mathcal{C}}(X_1,...,X_n) \xrightarrow{P} 0 \tag{2}$$

iff

$$\sup_{A \in \mathcal{C}} |P_n(A) - P(A)| \to 0 \text{ almost surely}. \tag{3}$$

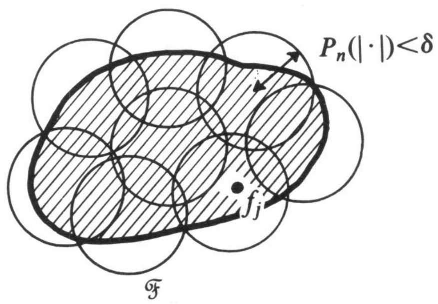
Measurability conditions are necessary because $\mathcal{C}$ might be uncountable. For instance the event $\Delta^{\mathcal{C}}(X_1) = 1$ occurs iff $X_1 \in (\bigcap_{\mathcal{C}} A) \cup (\bigcap_{\mathcal{C}} A^c)$. However, the intersection of uncountable many measurable sets need not be measurable. The same is true for the supremum over an uncountable $\mathcal{C}$. Thus, (2) and (3) do not always make sense. In Pollard's book, there is a thorough treatment of measurability issues.

Pollard's contributions to laws of large numbers are his methods of proof and the extension to classes $\mathcal{F}$ of, not necessarily uniformly bounded, measurable functions f. A collection $\mathcal{Q}$ of sets can be regarded as a special case by identifying $\mathcal{Q}$ with the class $\{1_A : A \in \mathcal{Q}\}$ of indicator functions. As with sets, the notation $P_n(f) = \int f dP_n$ and $P(f) = \int f dP$ is used. The distance for P_n between f and $\tilde{f}$ is defined as $P_n(|f-\tilde{f}|)$. Thus, for sets

$$P_n(|1_A - 1_{\tilde{A}}|) = P_n(A \Delta \tilde{A}).$$

Furthermore, for each $\delta > 0$, $N_1(\delta, P_n, \mathfrak{F})$ denotes the minimum number of balls with radius δ , necessary to cover $\mathfrak{F}$, i.e. $N_1(\delta, P_n, \mathfrak{F})$ is the smallest value of m such that there exist $f_1, ..., f_m$ such that for all $f \in \mathfrak{F}$,

$$\min_j P_n(|f - f_j|) < \delta.$$



The logarithm of $N_1(\delta, P_n, \mathfrak{F})$ is called the *δ-entropy* of $\mathfrak{F}$ for P_n. Observe that for a collection of sets $\Delta^{\mathcal{Q}}(X_1, ..., X_n) = N_1(\delta, P_n, \{1_A : A \in \mathcal{Q}\})$, $0 < \delta < 1/n$.

Pollard shows that if both

$$1^{-n} \log N_1(\delta, P_n, \mathfrak{F}) \xrightarrow{P} 0 \tag{4}$$

and

$$\int \sup_{f \in \mathfrak{F}} |f| dP < \infty \tag{5}$$

then

$$\sup_{f \in \mathfrak{F}} |P_n(f) - P(f)| \to 0 \text{ almost surely,} \tag{6}$$

again, provided certain measurability conditions are met.

The topics mentioned so far are all part of Chapter II. Let me now briefly present some of the results concerning weak convergence.

The uniform law of large numbers (6) can be formulated as the almost sure convergence of $P_n(\cdot)$ to $P(\cdot)$ as process in $f \in \mathfrak{F}$. That is, $P_n(\cdot)$, $n = 1,2,...$ and $P(\cdot)$ are considered as elements of some space $\mathfrak{X}$ of real-valued functions on $\mathfrak{F}$, $\mathfrak{X}$ being equipped with supremum norm. Note that the question whether or not $P_n(\cdot)$ is a *random* element of $\mathfrak{X}$ is not really relevant: basically only the measurability of the supremum of $|P_n(\cdot) - P(\cdot)|$ is of concern. However, for the study of weak convergence, $\nu_n(\cdot) := \sqrt{n}(P_n(\cdot) - P(\cdot))$ needs to be viewed as *random* element of $\mathcal{X}$.

Pollard explains the concept of weak convergence on general metric spaces and I shall sketch what it means for the process $\nu_n(\cdot)$. Some authors (e.g. DUDLEY and PHILIP [2]) circumvent the notion by creating large probability spaces where weak convergence is replaced by convergence in probability.

The supremum metric generally makes $\mathcal{X}$ into a nonseparable space. As a consequence, $\nu_n(\cdot)$ is not Borel-measurable. One already has to face this problem if $\mathcal{X}$ is the space D[0,1] of functions on [0,1] that are right-continuous and have left-hand limits. The classical solution is to take a different metric on D[0,1] (the Skorohod metric), but to stick to the Borel σ-algebra. Pollard eleborates on the alternative approach: maintain the supremum metric but choose a smaller σ-algebra for which $\nu_n(\cdot)$ is measurable. Under mild regularity conditions, $\nu_n(\cdot)$ will be measurable for the σ-algebra $\mathcal{B}^P$ that contains all closed balls with centres in a separable set in $\mathcal{X}$ and that makes all coordinate projections

$$\nu_n(f_1),\ldots,\nu_n(f_m)\ ;\ f_1,\ldots,f_m \in \mathcal{F}$$

measurable. In D[0,1], the σ-algebra generated by closed balls coincides with the σ-algebra that makes all coordinate projections measurable.

Remember that $\mathcal{X}$ is a space of functions y on $\mathcal{F}$. Such a y is continuous if $|y(f)-y(\tilde{f})| \to 0$ for all $f,\tilde{f}\in\mathcal{F}$, $P(|f-\tilde{f}|^2)\to 0$. A separable set in $\mathcal{X}$ is the collection of all bounded uniformly continuous functions. Now, the limiting distribution of $\nu_n(\cdot)$, if it exists, must be some Gaussian process on $\mathcal{F}$. An entropy condition on $\mathcal{F}$ ensures that there is a version of this Gaussian process with bounded, uniformly continuous sample paths. In other words, the limiting distribution concentrates on a separable set. This is important, because if one defines weak convergence with $\mathcal{B}^P$ as the σ-algebra on $\mathcal{X}$, and if the limiting distribution concentrates on a separable set, then some useful theorems for the Euclidean case (the continuous mapping theorem and the almost sure representation theorem) go through for weak convergence in more general $\mathcal{X}$.

Pollard's book also contains the equipment for verifying weak convergence. First, it runs into topics like relative compactness and uniform tightness of a collection of probability measures. A clear description of the *chaining method*, as a technique to prove tightness, is presented.

Several trips are made to other stochastic processes, apart from the empirical process $\nu_n(\cdot)$. Pollard succeeds in treating difficult topics in a transparent way. He sometimes jumps back to earlier sections of the book, to recall certain definitions or results. This helps the reader to keep track. The point of view he provides to classical situations is enlightning.

REFERENCES

1. P. BILLINGSLEY (1968). *Convergence of Probability Measures,* Wiley, New York.
2. R.M. DUDLEY, W. PHILIP (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verw. Geb. 62,* 509-552.

3. V.N. VAPNIK, Y.A. CHERVONENKIS (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Applications 16,* 264-280.

Abstracts

of Recent CWI Publications

When ordering any of the publications listed below please use the order form at the back of this issue.

CS-R8636. J.C. van Vliet & J.B. Warmer. *Intertable.*
AMS 68K05; CR I.7.2; 12 pp.; **key words:** document preparation.
Abstract: Intertable is an interactive program for editing tables. It can be used to create and manipulate tables. The present report describes Intertable, with emphasis on the algorithms used to incrementally update the widths of rows and columns of a table.

CS-R8637. P.J.W. ten Hagen, A.A.M. Kuijk & C.G. Trienekens. *Display architecture for VLSI-based graphics workstations.*
AMS 69K31, 69K33, 69K37; 15 pp.; **key words:** workstation architecture, computer graphics, interaction, raster, VLSI.
Abstract: At present, two popular development areas in computer graphics are improvement of interaction behaviour and more realistic graphics. The architecture for a high quality interactive workstation proposed in this work is designed such that both demanding and in a sense competing needs can be served. Calculations for generating realistic full 3-D scenes with lighting, transparency, reflection, and refraction effects, are done on the workstation itself. Intermediate results are stored to locally serve high level interaction mechanisms.

OS-R8614. G.A.P. Kindervater & J.K. Lenstra. *Parallel computing in combinatorial optimization.*
AMS 90C27, 68Q15, 68Q25, 68Rxx; 27 pp.; **key words:** parallel computer, computational complexity, polylog parallel algorithm, $\mathcal{P}$-completeness, sorting, shortest paths, minimum spanning tree, matching, maximum flow, linear

programming, knapsack, scheduling, traveling salesman, dynamic programming, branch and bound.

Abstract: This is a review of the literature on parallel computers and algorithms that is relevant for combinatorial optimization. We start by describing theoretical as well as realistic machine models for parallel computations. Next, we deal with the complexity theory for parallel computations and illustrate the resulting concepts by presenting a number of polylog parallel algorithms and $\mathcal{P}$-completeness results. Finally, we discuss the use of parallelism in enumerative methods.

OS-N8603. J.W. Polderman. *Adaptive pole assignment by state feedback.*
AMS 93C40; 8 pp.; **key words:** adaptive poleplacement, self-tuning, certainty-equivalence.

Abstract: An algorithm for adaptive poleplacement for a restricted class of systems is proposed. The asymptotic properties of the algorithm are analysed by studying the invariant points and the asymptotic active part of the state space. A weak form of self-tuning is derived.

NM-R8624. P.M. de Zeeuw. *NUMVEC FORTRAN library manual. Chapter: Elliptic PDEs. Routine: MGD1V and MGD5V.*
AMS 65V05, 65N20, 65F10; CR 5.17; 29 pp.; **key words:** elliptic PDEs, Galerkin approximation, multigrid methods, software, sparse linear systems, ILU relaxation, ILLU relaxation.

Abstract: The NUMVEC FORTRAN library routines MGD1V and MGD5V are described. These solve 7-diagonal linear systems arising from 7-point discretizations of elliptic PDEs on a rectangle, using a multigrid technique with ILU and ILLU relaxation respectively as smoothing process.

MS-R8613. R. Helmers & F.H. Ruymgaart. *Asymptotic normality of generalized L-statistics with unbounded scores.*
AMS 62G05, 62G30, 62E20; 8 pp.; **key words:** *U*-statistics, empirical processes, *L*-statistics, unbounded scores, asymptotic normality.

Abstract: A central limit theorem for linear combinations of a function of generalized order statistics with unbounded scores is established. The result supplements previous work of Silverman (1983), Serfling (1984) and Akritas (1986) concerning the asymptotic normality of generalized *L*-statistics. Our proof is patterned after the well-known Chernoff-Savage approach. A linear bound for the empirical distribution function of *U*-statistic structure is also derived and subsequently applied in the treatment of certain remainder terms.

MS-R8614. R.D. Gill & M.N. Voors (eds.). *Papers on semiparametric models at the ISI centenary session, Amsterdam.*
AMS 62G05, 62G10; 85 pp.; **key words:** semiparametric models, asymptotic efficiency, semiparametric Pareto model, Cox regression model, counting process, transformation models.

Abstract: This report contains revised versions of the papers presented at a meeting on semiparametric models during the ISI Centenary Session, together with the invited and open discussion, and one further paper. The three papers presented were: *Semiparametric models: progress and problems,* by J.A. Wellner; *The semiparametric Pareto model for regression analysis of survival times,* by D.G. Clayton and J. Cuzick; and *On asymptotic inference about intensity parameters of a counting process,* by K. Dzhaparidze. The extra paper by invited discussant P.J. Bickel, is *Efficient testing in a class of transformation models.*

PM-R8606. M. Hazewinkel. *Lie algebraic method in filtering and identification.*

AMS 93E11, 93B30, 93E10, 60H15, 93B15, 17B65, 17B99, 57R25; 17 pp.; **key words:** nonlinear filtering, estimation Lie algebra, asymptotic expansion, Weyl algebra, Heisenberg algebra, Kalman-Bucy filter, conditional density, Duncan-Mortensen-Zakai equation, BC-principle, identification, Lie algebra of vectorfields, finite dimensional filter, robustness.

Abstract: These lectures concern (nonlinear) filtering: very roughly, the art of obtaining best estimates for some stochastic time-varying variable x on the basis of observations of another process y. The more concrete object under consideration is a stochastic dynamical system $dx = f(x)dt + g(x)dw$, where w is Wiener noise, with observations $dy = h(x)dt + dv$, corrupted by further noise. The subject as presented here involves ideas and techniques from Lie algebra theory, stochastics, differential topology, approximation theory and partial differential equations and has relations with quantum theory and stochastic physics. The lectures are addressed to practitioners in any one of these areas assuming that as a rule they are not experts in the other ones.

PM-R8607. A.M. Cohen & G.M. Seitz. *The r-rank of the groups of exceptional Lie type.*

AMS 20G15, 20E15; 7 pp.; **key words:** groups of Lie type, elementary Abelian subgroups.

Abstract: In this note, we determine, for each simple group G of exceptional Lie type over an algebraically closed field F and each prime r distinct from the characteristic of F, the maximal number a for which there exists an elementary Abelian subgroup of order r^a in G. This settles a question raised in a paper of Borel & Serre (1953).

CWI Activities

Winter 1986

With each activity we mention its frequency and (between parentheses) a contact person at CWI. Sometimes some additional information is supplied, such as the location if the activity will not take place at CWI.

Study group on Analysis on Lie Groups. Jointly with University of Leiden. Biweekly. (T.H. Koornwinder)

International mini-conference on Lie Groups. Jointly with University of Leiden, Utrecht and Groningen. 21-22 April 1987. (T.H. Koornwinder)

Seminar on Integrable Systems. Once a month. (M. Hazewinkel)
A central object of study is the work of Belavin and Drinfeld, especially the relation between simple Lie algebras and solutions of the so-called classical Yang-Baxter equation. Also, linearization aspects of nonlinear representations and lattice KP, KdV will be discussed.

Seminar on Algebra and Geometry. Once a month. (A.M. Cohen)
The Cohomology of the Schubert variety and Coxeter groups.

Cryptography working group. Monthly. (J.H. Evertse)

Colloquium 'STZ' on System Theory, Applied and Pure Mathematics. Twice a month. (J. de Vries)

Study group 'Biomathematics'. Lectures by visitors or members of the group. Jointly with University of Leiden. Bimonthly (O. Diekmann)
Topics for the next meetings are: stochastic population dynamics, dynamics of structured populations.

Study group on Nonlinear Analysis. Lectures by visitors or members of the group. Jointly with University of Leiden. Bimonthly (O. Diekmann)
The purpose is to follow and investigate recent developments on qualitative analysis of nonlinear equations.

Progress meetings of the Applied Mathematics Department. Weekly (N.M.

Temme)

New results and open problems on the research topics of the department: biomathematics, mathematical physics, asymptotic and applied analysis, image analysis.

Study group on Statistical and Mathematical Image Analysis. Every three weeks. (R.D. Gill)

The group is presently studying J. Serra's approach to image analysis, 'mathematical morphology', and recent statistical contributions using Markov field modelling due to S. and D. Geman, J. Besag and B. Ripley.

Progress meetings of the Mathematical Statistics Department. Biweekly (H.C.P. Berbee)

Talks by members of the department on recent developments in research and consultation.

Study group on Empirical Processes. Jointly with University of Amsterdam. Biweekly (S. van de Geer)

The group is studying the recent book *Convergence of Stochastic Processes* by D. Pollard, and related literature.

System Theory Days. Irregular. (J.H. van Schuppen, J.M. Schumacher)

Study group on System Theory. Biweekly. (J.M. Schumacher)

Current topic: Discrete event dynamical systems.

Colloquium on Queueing Theory and Performance Evaluation. Irregular. (O.J. Boxma)

Progress meetings on Numerical Mathematics. Weekly. (H.J.J. te Riele)

Study group on Numerical Software for Vector Computers. Monthly. (H.J.J. te Riele)

Study group on Differential and Integral Equations. Lectures by visitors or group members. Irregular. (H.J.J. te Riele)

Study group on Graphics Standards. Monthly. (M. Bakker)

Study group on Dialogue Programming. (P.J.W. ten Hagen)

Process Algebra Meeting. Weekly. (J.W. Klop)

Visitors to CWI from Abroad

E. Badertscher (University of Bern, Switzerland) 16-19 December. J.A. Ball (Virginia Tech., Blacksburg, USA) 26 November. L. Baratchart (INRIA, Sophia Antipolis, Valbonne, France) 26-28 November. R.S. Bird (Oxford University, England) 10-18 December. R.K. Boel (University of Ghent, Belgium) 18 November. H. Bruneel (University of Ghent, Belgium) 15 December. H. Brunner (Memorial University of Newfoundland, St. John, Canada) 15-17 October. B.A. Francis (University of Toronto, Canada) 13-17 December. A. Jakubowicz (University of Szczecin, Poland) 15-24 October. K.E. Karlsson (ASEA Research and Innovation, Vasteras, Sweden) 8 December. D.A. Leites (Stockholm University, Sweden) 16 November - 6 December. Y. Moses (Weizman Institute, Rehovot, Israel) 15-16 December. J.M. Sanz-Serna (University of Valladolid, Spain) 1-4 October. F. Soumis (Ecole Polytechnical, Montréal, Canada) 15-17 December. B. Zwahler (EPFL, Lausanne, Switzerland) 9 December.

Order Form for CWI Publications

Sales Department
Centre for Mathematics and Computer Science
Kruislaan 413
1098 SJ Amsterdam
The Netherlands

☐ Please send the publications marked below on an exchange basis
☐ Please send the publications marked below with an invoice

	Publication code	Price per copy	Number of copies wanted
☐	CS-R8636	4.--	....
☐	CS-R8637	4.--	....
☐	OS-R8614	5.20	....
☐	OS-N8603	4.--	....
☐	NM-R8624	5.20	....
☐	MS-R8613	4.--	....
☐	MS-R8614	12.70	....
☐	PM-R8606	4.--	....
☐	PM-R8607	4.--	....

If you wish to order any of the above publications please tick the appropriate boxes and return the completed form to our Sales Department.
Don't forget to add your name and address!
Prices are given in Dutch guilders and are subject to change without notice. Foreign payments are subject to a surcharge per remittance to cover bank, postal and handling charges.

Name ....

Street ....

City ....

Country ....

Signature ....

Date ....

Contents