Werk

Titel: Journal für die reine und angewandte Mathematik
Verlag: de Gruyter
Jahr: 1986
Kollektion: Mathematica
Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen
Werk Id: PPN243919689_0365
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN243919689_0365

Übergeordnetes Werk

Werk Id: PPN243919689
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN243919689

Terms and Conditions

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

Journal für die reine und angewandte Mathematik

gegründet 1826 von

August Leopold Crelle

fortgeführt von
C. W. Borchardt, K. Weierstrass, L. Kronecker, L. Fuchs,
K. Hensel, L. Schlesinger, H. Hasse, H. Rohrbach

gegenwärtig herausgegeben von

Willi Jäger · Martin Kneser · Horst Leptin
Samuel J. Patterson · Peter Roquette · Michael Schneider

unter Mitwirkung von
J. Arthur (Toronto), T. tom Dieck (Göttingen), O. Forster (München),
P. R. Halmos (Bloomington), F. Hirzebruch (Bonn),
R. Howe (New Haven), Y. Ihara (Tokyo), H. Koch (Berlin), J. Lindenstrauss (Jerusalem)

JRMAA8

Band 365

W DE G

Walter de Gruyter · Berlin · New York 1986

Printed in Germany. Satz und Druck: Arthur Collignon GmbH, 1000 Berlin 30

ISSN 0075-4102

Inhalt

Indexed in Science Citation Index, ASCA, and ISI/COMPUMATH.
Covered by Zentralblatt für Mathematik/Mathematics Abstract.

Ausgabedatum des Bandes 365

14. März 1986

Products of distributions

By *M. Oberguggenberger* at Innsbruck

0. Introduction

The purpose of this article is to simplify the definition of the multiplicative product of two distributions given by Ambrose [1] in this Journal and to establish the relationship between:

(a) this definition;

(b) Mikusiński's product [6] (which has been shown by Shiraishi-Itano [11] to be equivalent to a product of Hirata-Ogata [3]);

(c) a localized version of Vladimirov's approach [14] using the exchange formula;

(d) a modified version of the definition given by Reed-Simon [8];

(e) Kamiński's Δ-product [5];

(f) multiplication in Colombeau's algebra $\mathcal{G}(\mathbb{R}^n)$.

Let U and V be members of $\mathcal{D}'(\mathbb{R}^n)$.

Definition 0. 1. The multiplicative product of U and V *in the sense of Mikusiński-Hirata-Ogata*, denoted by $[U]\,[V]$, is said to exist if the following condition holds:

(MHO) For all δ-sequences $\{\rho_j, j \in \mathbb{N}\}$, $\{\sigma_j, j \in \mathbb{N}\}$ the limit $\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V)$ exists in $\mathcal{D}'(\mathbb{R}^n)$.

In this case the limit is independent of the δ-sequence chosen, and $[U]\,[V]$ is defined to be this limit.

Here $\{\rho_j, j \in \mathbb{N}\} \subset \mathcal{D}(\mathbb{R}^n)$ is called a *δ-sequence*, if for all $j \in \mathbb{N}$

(δ 1) there are positive numbers $\varepsilon_j \to 0$ so that $\rho_j(x) = 0$ if $|x| \geqq \varepsilon_j$,

(δ 2) $\int \rho_j(x)\, dx = 1$,

(δ 3) $\rho_j \geqq 0$.

Definition 0. 2. The multiplicative product of U and V *in the sense of Ambrose*, denoted by $U \cdot V$, is said to exist, if for every $x \in \mathbb{R}^n$ there exists a neighborhood Q_x such that for all $\omega, \psi \in \mathcal{D}(Q_x)$

(A 1) $\mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V)$ belongs to $L^1(\mathbb{R}^n)$,

(A 2) $\int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V) = \int \mathcal{F}(\omega V)\,\mathcal{F}^{-1}(\psi U)$,

(A 3) the map $\omega \mapsto \int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V) : \mathcal{D}(Q_x) \to \mathbb{C}$ is continuous.

The expression $\int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V)$, where $\psi \in \mathcal{D}(Q_x)$ is chosen so that $\psi \equiv 1$ on the support of ω, defines a continuous linear form for $\omega \in \mathcal{D}(Q_x)$, once the independence of ψ is verified. $U \cdot V$ is defined to be the unique distribution on $\mathbb{R}^n$ coinciding with these local distributions on each Q_x.

Here $\mathcal{F}$ denotes the Fourier transform which for $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is given by

$$\mathcal{F}\varphi(\xi) = \int \exp(-2\pi i x\xi)\,\varphi(x)\,dx.$$

So far it has been shown by Tysk [13] (see also Colombeau [2], § 2. 3) that conditions (A 1) (A 2) (A 3) together imply (MHO) and that under these circumstances $U \cdot V$ equals $[U]\,[V]$. Here we show that already (A 1) implies (MHO) and, moreover, (A 1) implies (A 2) and (A 3). We employ a criterion for the existence of $[U]\,[V]$ due to Shiraishi-Itano [11], thereby avoiding Tysk's calculations with δ-sequences. As a result, the proof that Ambrose' order condition defines a "domain of multiplication" is now much shorter. Namely, it is checked at once that U and V satisfy (A 1), provided for every $(x, \xi) \in T^*\mathbb{R}^n \setminus \{0\}$ there is $s, t \in \mathbb{R}$ with $s + t \geqq 0$ so that microlocally $U \in H^s(x, \xi)$ and $V \in H^t(x, -\xi)$, see Section 3.

Recall that the $\mathcal{S}'$-convolution of two tempered distributions S, T is said to exist if $(\varphi * \tilde{S})\,T$ belongs to $\mathcal{D}'_{L^1}$ for every $\varphi \in \mathcal{S}$. In this case $S * T$ is defined by $\varphi \mapsto \langle (\varphi * \tilde{S})\,T, 1\rangle$ and belongs to $\mathcal{S}'$, see Shiraishi [9]; $\tilde{S}$ denotes the homothetic image of S under multiplication by (-1). The remark in Vladimirov [14], Section 6. 5, p. 114 suggests to define the product of two tempered distributions U, V by $\mathcal{F}^{-1}(\mathcal{F}U * \mathcal{F}V)$ provided the $\mathcal{S}'$-convolution of $\mathcal{F}U$ and $\mathcal{F}V$ exists. We localize this approach as follows:

Definition 0. 3. Two distributions $U, V \in \mathcal{D}'(\mathbb{R}^n)$ are said to satisfy the *localized Vladimirov condition* if the following holds:

(LV) For every $x \in \mathbb{R}^n$ there is a neighborhood Ω_x and $f \in \mathcal{D}(\mathbb{R}^n)$ with $f \equiv 1$ on Ω_x so that the $\mathcal{S}'$-convolution of $\mathcal{F}(fU)$ and $\mathcal{F}(fV)$ exists.

Using (LV) a new product $U \circ V$ may be defined (see Section 2). We show, however, that (LV) is equivalent to (A 1) and that the products $U \circ V$ and $U \cdot V$ coincide, if they exist. As a consequence, if U and V are tempered distributions such that the $\mathcal{S}'$-convolution of $\mathcal{F}U$ and $\mathcal{F}V$ exists, then $U \cdot V$ exists and equals $\mathcal{F}^{-1}(\mathcal{F}U * \mathcal{F}V)$. This sharpens the exchange formula of Hirata-Ogata [3]. Also, a modified version of the definition of Reed-Simon [8] will be shown to imply (LV).

Finally, to put the Mikusiński-Hirata-Ogata product into relation with Kamiński's multiplication [5] and Colombeau's algebras [2], we need to get rid of the positivity requirement ($\delta 3$) on the δ-sequences employed. The appropriate trick is provided in Section 3. Consequently, the existence of $[U]\,[V]$ implies the existence of the Δ-product of U and V; the latter implies that the product of U and V in Colombeau's algebra $\mathscr{G}(\mathbb{R}^n)$ admits an "associated distribution", and all three notions are consistent.

Concerning our notation we follow Horváth [4], except that the space of integrable distributions will be denoted by $\mathscr{D}'_{L^1}(\mathbb{R}^n)$. For the sequential definitions of the product we employ Kamiński's suggestive notation using square brackets. The reader should be warned that square brackets have a different meaning in [1] and [11].

1. On the product of Ambrose

The main tool in this section will be the following criterion for the existence of $[U]\,[V]$ which is due to Shiraishi-Itano.

Proposition 1. 1. *For $U, V \in \mathscr{D}'(\mathbb{R}^n)$ each of the following conditions is equivalent to* (MHO):

(a) *For all δ-sequences $\{\rho_j, j \in \mathbb{N}\}$ (satisfying ($\delta 1$), ($\delta 2$), ($\delta 3$)) the limit $\lim_{j\to\infty} (\rho_j * U)V$ exists in $\mathscr{D}'(\mathbb{R}^n)$.*

(b) *For all δ-sequences $\{\sigma_j, j \in \mathbb{N}\}$ (satisfying ($\delta 1$), ($\delta 2$), ($\delta 3$)) the limit $\lim_{j\to\infty} U(\sigma_j * V)$ exists in $\mathscr{D}'(\mathbb{R}^n)$.*

(c) *For all $\chi \in \mathscr{D}(\mathbb{R}^n)$ there is a neighborhood Ω of zero so that $(\chi U) * \check{V}$ belongs to $L^\infty(\Omega)$ and is continuous at 0.*

If the limits in (a) *and* (b) *exist, then they are independent of the δ-sequences chosen and are denoted by $[U]\,V$ and $U[V]$, respectively. Moreover, if U, V satisfy any of the equivalent conditions* (a), (b), (c), (MHO), *then*

$$\langle [U]\,[V], \chi\rangle = \langle [U]\,V, \chi\rangle = \langle U[V], \chi\rangle = ((\chi U) * \check{V})\,(0)$$

for $\chi \in \mathscr{D}(\mathbb{R}^n)$.

Proof. The assertions follow from Propositions 1, 2, 3, and 5 of Shiraishi-Itano [11]. □

Theorem 1. 2. *Suppose $U, V \in \mathscr{D}'(\mathbb{R}^n)$ satisfy condition* (A1). *Then U, V satisfy* (A2) *and* (A3) *as well, thus $U \cdot V$ exists. Moreover, $[U]\,[V]$ exists and is equal to $U \cdot V$.*

Proof. Step 1. (A1) *implies* (A3): Let $x \in \mathbb{R}^n$ and Q_x be the corresponding neighborhood where (A1) holds. Consider the composition of linear maps

$$L : \mathscr{D}(Q_x) \longrightarrow \mathscr{E}' \longrightarrow \mathscr{S}' \longrightarrow \mathscr{S}',$$

$$\omega \longmapsto \omega U \longmapsto \mathscr{F}(\omega U) \longmapsto \mathscr{F}(\omega U)\,\mathscr{F}^{-1}(\psi V)$$

for fixed $\psi \in \mathscr{D}(Q_x)$. The first two maps obviously are continuous, the last map is continuous because $\mathscr{F}^{-1}(\psi V) \in O_M$ and multiplication by a fixed element of O_M is a continuous map from $\mathscr{S}'$ to $\mathscr{S}'$. Thus $L : \mathscr{D}(Q_x) \to \mathscr{S}'$ is continuous. By (A1), the range of L is contained in $L^1(\mathbb{R}^n)$. The closed graph theorem (in the version of [4], Theorem 3. 17. 4) now implies the continuity of $L : \mathscr{D}(Q_x) \to L^1(\mathbb{R}^n)$, whence (A3).

Step 2. *For any* $\omega \in \mathcal{D}(Q_x)$, $[\omega U]\,[V]$ *exists, and for any* $\psi \in \mathcal{D}(Q_x)$ *with* $\psi \equiv 1$ *on* $\operatorname{supp}\omega$ and any $\chi \in \mathcal{D}(\mathbb{R}^n)$,

$$\langle [\omega U]\,[V], \chi\rangle = \int \mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V). \tag{1}$$

To prove this, let first ω, ψ be arbitrary members of $\mathcal{D}(Q_x)$, $\chi \in \mathcal{D}(\mathbb{R}^n)$. By (A1), $\mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V)$ belongs to $L^1(\mathbb{R}^n)$, thus $\mathcal{F}^{-1}\big(\mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V)\big)$ is a continuous function on $\mathbb{R}^n$. Since $\chi\omega U \in O_C'$ and $\psi V \in \mathcal{S}'$ we have by the exchange theorem (see [4], Theorem 4. 11. 3.)

$$\mathcal{F}^{-1}\big(\mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V)\big) = (\chi\omega U) * (\psi V)^{\sim}.$$

By criterion (c) of Proposition 1. 1, this means that ωU and ψV satisfy (MHO), and

$$\langle [\omega U]\,[\psi V], \chi\rangle = \big((\chi\omega U) * (\psi V)^{\sim}\big)(0) = \int \mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V). \tag{2}$$

Take now a particular $\varphi \in \mathcal{D}(Q_x)$ which is identically one in a neighborhood Q of $\operatorname{supp}\omega$. Then for any δ-sequence $\{\rho_j, j \in \mathbb{N}\}$, $\lim_{j\to\infty} (\omega U * \rho_j)(\varphi V)$ exists by criterion (a) of Proposition 1. 1. For large j, $\operatorname{supp}(\omega U * \rho_j) \subset Q$, so

$$(\omega U * \rho_j)(\varphi V) = (\omega U * \rho_j)\, V.$$

Together with Proposition 1. 1 (a) this proves the existence of $[\omega U]\,[V]$.

By the Theorem in [11], p. 227, the existence of $[\omega U]\,V$ implies the existence of $[\omega U]\,[\psi V]$ for arbitrary $\psi \in \mathcal{E}(\mathbb{R}^n)$ and the equality of $[\omega U]\,[\psi V]$ with $[\omega\psi U]\,[V]$. Thus if $\psi \equiv 1$ on $\operatorname{supp}\omega$, and $\chi \in \mathcal{D}(\mathbb{R}^n)$, then

$$\langle [\omega U]\,[V], \chi\rangle = \langle [\omega U]\,[\psi V], \chi\rangle = \int \mathcal{F}(\chi\omega U)\,\mathcal{F}^{-1}(\psi V).$$

Step 3. *The expression* $\int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V)$ *is independent of* $\psi \in \mathcal{D}(Q_x)$ *as long as* $\psi \equiv 1$ *on* $\operatorname{supp}\omega$. Indeed, taking $\chi \in \mathcal{D}(\mathbb{R}^n)$ identically one on Ω_x, we have by (1)

$$\int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V) = \langle [\omega U]\,[V], \chi\rangle$$

which is independent of ψ. Step 1 now shows that for every x the map

$$\omega \longmapsto \int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V)$$

(with $\psi \equiv 1$ on $\operatorname{supp}\omega$) defines a distribution on Q_x. There is a unique $W \in \mathcal{D}'(\mathbb{R}^n)$ which coincides with these local distributions on each Q_x. If we let $\{\omega_k, k \in \mathbb{N}\}$ be a locally finite, smooth partition of unity subordinate to the cover $\{Q_x : x \in \mathbb{R}^n\}$ with $\operatorname{supp}\omega_k \subset Q_{x_k}$, and $\psi_k \in \mathcal{D}(Q_{x_k})$ with $\psi_k \equiv 1$ on $\operatorname{supp}\omega_k$, then W is given by

$$\langle W, \chi\rangle = \sum_{k=1}^{\infty} \int \mathcal{F}(\chi\omega_k U)\,\mathcal{F}^{-1}(\psi_k V).$$

Step 4. $[U]\,[V]$ *exists and equals* W: Take $\chi \in \mathcal{D}(\mathbb{R}^n)$ and let $\{\rho_j, j \in \mathbb{N}\}$ be a δ-sequence. Then

$$\lim_{j\to\infty} \langle U(V * \rho_j), \chi\rangle = \sum_{k=1}^{\infty} \lim_{j\to\infty} \langle \omega_k U(V * \rho_j), \chi\rangle.$$

Since $\omega_k \in \mathcal{D}(Q_{x_k})$, Step 2 implies that

$$\lim_{j\to\infty} \langle \omega_k U(V * \rho_j), \chi\rangle = \langle [\omega_k U]\,[V], \chi\rangle = \int \mathcal{F}(\chi\omega_k U)\,\mathcal{F}^{-1}(\psi_k V).$$

Thus $[U]\,[V] = U[V] = \lim_{j\to\infty} U(V * \rho_j)$ exists and equals W.

Step 5. *U, V satisfy* (A2): Let $x \in \mathbb{R}^n$ and ω, ψ be arbitrary members of $\mathcal{D}(Q_x)$. By the Theorem in [11], p. 227, the existence of $[U]\,[V]$ implies that $[\omega U]\,[\psi V]$ exists and equals $[\omega V]\,[\psi U]$. Taking $\chi \in \mathcal{D}(\mathbb{R}^n)$ identically one on Q_x, we have by (2)

$$\int \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V) = \langle [\omega U]\,[\psi V], \chi\rangle = \langle [\omega V]\,[\psi U], \chi\rangle = \int \mathcal{F}(\omega V)\,\mathcal{F}^{-1}(\psi U).$$

That V and U (reversed order) satisfy (A1), if U and V do, follows by a change of variable in the integrand; thus the last equality is a consequence of (2) again. We conclude that (A2) holds, so we are justified to say that $U \cdot V$ exists in the sense of Ambrose; obviously $U \cdot V = W = [U]\,[V]$. □

Remarks. (a) The proof of Step 1 shows that the bilinear map

$$\mathcal{D}(Q_x) \times \mathcal{D}(Q_x) \longrightarrow L^1(\mathbb{R}^n) : (\omega, \psi) \longmapsto \mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V)$$

is separately continuous; since $\mathcal{D}(Q_x)$ is barrelled, it is also sequentially continuous. This provides a short proof of Lemma 3. 2 of [1].

(b) The existence of $[U]\,[V]$ does not imply the existence of $U \cdot V$. For instance, if $U \cdot \delta$ exists, then there is a neighborhood Q_0 of zero so that

$$\mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi \delta) = \psi(0)\,\mathcal{F}(\omega U) \in L^1(\mathbb{R}^n)$$

for all $\omega, \psi \in \mathcal{D}(Q_0)$, thus U is continuous in a neighborhood of zero. On the other hand, $[U]\,[\delta]$ exists if and only if $\chi\delta * \check{U} = \chi(0)\,\check{U}$ is bounded on some neighborhood of zero and continuous at zero for every $\chi \in \mathcal{D}(\mathbb{R}^n)$. Thus if we take $U \in L^\infty(\mathbb{R}^n)$ which is continuous at 0 but not continuous in any neighborhood of 0, then $[U]\,[\delta]$ exists but $U \cdot \delta$ does not exist.

2. On the exchange formula

For the $\mathcal{S}'$-convolution of distributions the following associative law holds, which is due to Hirata-Ogata. This is Proposition 1 of [3]:

Lemma 2. 1. *Let $S, T \in \mathcal{S}'$ so that the $\mathcal{S}'$-convolution of S and T exists. If $R \in \mathcal{O}_C'$ then the $\mathcal{S}'$-convolutions of R and $(S * T)$ and of $(R * S)$ and T exist and $R * (S * T) = (R * S) * T$.* □

Take two tempered distributions U, V that satisfy the localized Vladimirov condition (LV). We are going to show how U and V may be multiplied. Let $x, y \in \mathbb{R}^n$ with corresponding test functions f, g identically one on neighborhoods Ω_x, Ω_y, as required by condition (LV). By assumption, the map

$$(3) \qquad \chi \longmapsto \langle \mathcal{F}^{-1}\chi, \mathcal{F}(fU) * \mathcal{F}(fV)\rangle \quad \text{for} \quad \chi \in \mathcal{D}(\Omega_x)$$

defines an element of $\mathcal{D}'(\Omega_x)$. We now apply Lemma 2. 1, the commutativity of the $\mathcal{S}'$-convolution, the exchange formula for the convolution on $\mathcal{S} \times \mathcal{S}'$, and the formula

$$g^2 = \mathcal{F}^{-1}(\mathcal{F}g * \mathcal{F}g)$$

to obtain

$$g^2\,\mathcal{F}^{-1}\big(\mathcal{F}(fU) * \mathcal{F}(fV)\big) = \mathcal{F}^{-1}\big(\mathcal{F}(gfU) * \mathcal{F}(gfV)\big) = f^2\,\mathcal{F}^{-1}\big(\mathcal{F}(gU) * \mathcal{F}(gV)\big).$$

Thus, if $\chi \in \mathcal{D}(\Omega_x \cap \Omega_y)$, then $\chi = \chi f^2 = \chi g^2$, and so

$$\langle \chi, \mathcal{F}^{-1}(\mathcal{F}(fU) * \mathcal{F}(fV))\rangle = \langle \chi, \mathcal{F}^{-1}(\mathcal{F}(gU) * \mathcal{F}(gV))\rangle.$$

This shows that there is a unique element of $\mathcal{D}'(\mathbb{R}^n)$ which coincides with the local distributions given by (3) on each Ω_x; this defines our product $U \circ V$.

Theorem 2. 2. *Let $U, V \in \mathcal{D}'(\mathbb{R}^n)$. Then U, V satisfy* (LV) *if and only if they satisfy* (A 1). *Moreover, the products $U \cdot V$ and $U \circ V$ coincide when they exist.*

Proof. Step 1. (LV) $\Rightarrow$ (A 1). Suppose U, V satisfy (LV). Let $x \in \mathbb{R}^n$, Ω_x a neighborhood of x and $f \in \mathcal{D}(\mathbb{R}^n)$ with $f \equiv 1$ on Ω_x as in condition (LV). By Theorem 3 of Shiraishi [9], the $\mathcal{S}'$-convolution of $\mathcal{F}(fU)$ and $\mathcal{F}(fV)$ exists if and only if

$$(\alpha * \mathcal{F}(fU))(\beta * \mathcal{F}^{-1}(fV)) \in L^1(\mathbb{R}^n)$$

for all $\alpha, \beta \in \mathcal{S}$. In particular, taking $\alpha = \mathcal{F}\omega$ and $\beta = \mathcal{F}\psi$ with $\omega, \psi \in \mathcal{D}(\Omega_x)$ and noting that $f \equiv 1$ on $\operatorname{supp}\omega$ and $\operatorname{supp}\psi$ we conclude that

$$\mathcal{F}(\omega U)\,\mathcal{F}^{-1}(\psi V) \in L^1(\mathbb{R}^n).$$

Thus (A 1) holds.

Step 2. (A 1) $\Rightarrow$ (LV). Suppose that U, V satisfy (A 1). Let $x \in \mathbb{R}^n$ and Q_x be a neighborhood as in condition (A 1). Pick another neighborhood $\Omega_x \Subset Q_x$ and $f \in \mathcal{D}(Q_x)$ with $f \equiv 1$ on Ω_x. We show that the $\mathcal{S}'$-convolution of $\mathcal{F}(fU)$ and $\mathcal{F}(fV)$ exists. Indeed, take $\alpha \in \mathcal{S}$. Then condition (A 1) implies that

$$\mathcal{F}(fU)\,(\alpha * \mathcal{F}^{-1}(fV) = \mathcal{F}(fU)\,\mathcal{F}^{-1}((\mathcal{F}\alpha) fV)$$

belongs to $L^1(\mathbb{R}^n)$. Since $L^1 \subset \mathcal{D}'_{L^1}$, U and V satisfy condition (LV).

Step 3. *Consistency.* Let Q_x, Ω_x and f be as in Step 2. For every $\omega \in \mathcal{D}(\Omega_x)$ we have

$$\begin{aligned}\langle U \cdot V, \omega\rangle &= \int \mathcal{F}(\omega f U)\,\mathcal{F}^{-1}(fV) = \int \mathcal{F}^{-1}(\omega f U)\,\mathcal{F}(fV)\\ &= \langle (\mathcal{F}^{-1}\omega * \mathcal{F}^{-1}(fU))\,\mathcal{F}(fV), 1\rangle = \langle \mathcal{F}(fU) * \mathcal{F}(fV), \mathcal{F}^{-1}\omega\rangle\\ &= \langle U \circ V, \omega\rangle.\end{aligned}$$

The first equality follows from Step 3 in the proof of Theorem 1. 2, since $f \in \mathcal{D}(Q_x)$ and $f \equiv 1$ on $\operatorname{supp}\omega$; the second equality is a change of variable in the integrand, the fourth equality follows from Step 2 where the existence of the $\mathcal{S}'$-convolution of $\mathcal{F}(fU)$ and $\mathcal{F}(fV)$ has been proven. We conclude that $U \cdot V = U \circ V$. □

Proposition 2. 3. *Let $U, V \in \mathcal{S}'(\mathbb{R}^n)$ such that the $\mathcal{S}'$-convolution of $\mathcal{F}U$ and $\mathcal{F}V$ exists. Then U and V satisfy the equivalent conditions* (A 1) *and* (LV), *thus $U \cdot V$ exists, and*

$$U \cdot V = \mathcal{F}^{-1}(\mathcal{F}U * \mathcal{F}V).$$

Proof, Given $f \in \mathcal{D}(\mathbb{R}^n)$, a repeated application of Lemma 2. 1 shows that the $\mathcal{S}'$-convolution of $\mathcal{F}(fU)$ and $\mathcal{F}(fV)$ exists. Thus U, V satisfy (LV). Let $x \in \mathbb{R}^n$ and take Q_x, Ω_x, and $f \in \mathcal{D}(Q_x)$ with $f \equiv 1$ on Ω_x as in Step 2 of the proof of Theorem 2. 2. If $\omega \in \mathcal{D}(\Omega_x)$, then

$$\begin{aligned}\langle \omega, \mathcal{F}^{-1}(\mathcal{F}U * \mathcal{F}V)\rangle &= \langle \omega, f^2\,\mathcal{F}^{-1}(\mathcal{F}U * \mathcal{F}V)\rangle\\ &= \langle \omega, \mathcal{F}^{-1}(\mathcal{F}(fU) * \mathcal{F}(fV))\rangle = \langle \omega, U \circ V\rangle\\ &= \langle \omega, U \cdot V\rangle.\end{aligned}$$

□

Remarks. (a) Hirata-Ogata [3] had shown that the existence of the $\mathscr{S}'$-convolution of $\mathscr{F}U$ and $\mathscr{F}V$ implies the existence of the product $[U][V]$ and the validity of the exchange formula. Our conclusion is that even $U \cdot V$ exists and satisfies the exchange formula, which is a stronger result in view of the remark at the end of Section 1.

(b) The exchange formula given by Proposition 2. 3 is not symmetric: There are tempered distributions U, V so that $U \cdot V$ exists and is tempered, but the $\mathscr{S}'$-convolution of $\mathscr{F}U$ and $\mathscr{F}V$ does not exist. In fact, $U = \delta(x)$ and $V = \delta(x-a)$ with $a \neq 0$ will do, since $(\varphi * 1)\exp(-2\pi i a\xi) \notin \mathscr{D}'_{L^1}$ if $\varphi \in \mathscr{S}$ is an even function, as follows from Proposition 8 of Ortner-Wagner [7], p. 363. This example has been communicated to the author by N. Ortner.

3. Applications and other products

As an immediate consequence of Theorem 1. 2 we can derive a microlocal condition ensuring the existence of $U \cdot V$. This condition is more or less a reformulation of Ambrose' order condition in terms of the H^s language; it contains Hörmander's wave front set criterion. To be specific, if $U \in \mathscr{D}'(\mathbb{R}^n)$, $s \in \mathbb{R}$, and $(x, \xi) \in T^*\mathbb{R}^n \setminus \{0\}$, then U is said to belong to $H^s(x, \xi)$, if $U = U_1 + U_2$ with $U_1 \in H^s(\mathbb{R}^n)$ and $(x, \xi) \notin WF(U_2)$, where $WF(.)$ denotes the wave front set; see for example Taylor [12], Section VI. 1.

Corollary 3. 1. *Let $U, V \in \mathscr{D}'(\mathbb{R}^n)$ and suppose that either*

(a) *for each $(x, \xi) \in T^*\mathbb{R}^n \setminus \{0\}$ there is $s = s(x, \xi)$ and $t = t(x, \xi) \in \mathbb{R}$ with $s + t \geqq 0$ so that $U \in H^s(x, \xi)$ and $V \in H^t(x, -\xi)$, or*

(b) $WF(U) \oplus WF(V) \equiv \{(x, \xi_1 + \xi_2) : (x, \xi_1) \in WF(U), (x, \xi_2) \in WF(V)\}$ *does not contain an element of the form $(x, 0)$.*

Then U, V satisfy (A1), *thus $U \cdot V$ exists.*

Proof. Suppose first that U, V satisfy (a). Fix (x, ξ), write $U = U_1 + U_2$ with $U_1 \in H^s(\mathbb{R}^n)$ and $(x, \xi) \notin WF(U_2)$, and let Q be a neighborhood of x, $f \in \mathscr{D}(\mathbb{R}^n)$ with $f \equiv 1$ on Q so that $\mathscr{F}(fU_2)$ is rapidly decreasing in some conical neighborhood Γ of ξ. An elementary calculation using Peetre's inequality shows that

$$\mathscr{F}(\omega U_2) = \mathscr{F}\omega * \mathscr{F}(fU_2)$$

is rapidly decreasing in some smaller cone $\Gamma' \subset \Gamma$ for every $\omega \in \mathscr{D}(Q)$. Thus

$$(1 + |.|^2)^{\frac{s}{2}} \mathscr{F}(\omega U) \in L^2(\Gamma')$$

for every $\omega \in \mathscr{D}(Q)$. Since $V \in H^t(x, -\xi)$ we have (possibly after shrinking Q and Γ') that

$$(1 + |.|^2)^{\frac{t}{2}} \mathscr{F}^{-1}(\psi V) \in L^2(\Gamma')$$

for every $\psi \in \mathscr{D}(Q)$. Since $s + t \geqq 0$, Hölder's inequality implies that $\mathscr{F}(\omega U)\mathscr{F}^{-1}(\psi V)$ belongs to $L^1(\Gamma')$. Since every covering of $\mathbb{R}^n \setminus \{0\}$ by open cones has a finite subcovering, U and V are seen to satisfy (A1).

That (b) implies (a) is obvious. □

Remark. The Hölder's inequality argument is the same as in the first part of Ambrose' proof of his Theorem 3. 1 [1], p. 80/81. Our proof of existence of $U \cdot V$ is simpler in as much as we no longer need to show that (A 2) and (A 3) hold.

As an application of Theorem 2. 2 we can give another condition ensuring the existence of $U \cdot V$. This is essentially the definition of the product of U and V given by Reed-Simon in [8], Section IX. 10:

Corollary 3. 2. *Let $U, V \in \mathscr{D}'(\mathbb{R}^n)$ and suppose that for every $x \in \mathbb{R}^n$ there is a neighborhood Ω_x and $f \in \mathscr{D}(\mathbb{R}^n)$ with $f \equiv 1$ on Ω_x so that for almost all $\xi \in \mathbb{R}^n$*

$$\int |\mathscr{F}(fU)(\eta)\, \mathscr{F}(fV)(\xi - \eta)|\, d\eta < \infty$$

and the map

$$\xi \longmapsto \int |\mathscr{F}(fU)(\eta)\, \mathscr{F}(fV)(\xi - \eta)|\, d\eta$$

is polynomially bounded (that is, equal to a polynomial times an L^∞-function). Then U and V satisfy (LV), *thus $U \cdot V$ exists.*

Proof. We show that the $\mathscr{S}'$-convolution of $\mathscr{F}(fU)$ and $\mathscr{F}(fV)$ exists. Let $\varphi \in \mathscr{S}$. Since $\mathscr{F}(fU)$ itself is polynomially bounded,

$$(\varphi * \mathscr{F}(fU)^\sim)(\xi) = \int \varphi(\eta)\, \mathscr{F}(fU)(\eta - \xi)\, d\eta.$$

By Fubini's theorem,

$$\begin{aligned} &\int |(\varphi * \mathscr{F}(fU)^\sim)(\xi)\, \mathscr{F}(fV)(\xi)|\, d\xi \\ &\quad \leqq \int |\varphi(\eta)| \int |\mathscr{F}(fU)(\eta - \xi)\, \mathscr{F}(fV)(\xi)|\, d\xi\, d\eta \end{aligned}$$

which is finite by hypothesis. Thus $(\varphi * \mathscr{F}(fU)^\sim)\, \mathscr{F}(fV) \in L^1 \subset \mathscr{D}'_{L^1}$. □

To determine the relation of the Mikusiński-Hirata-Ogata product with Kamiński's "Δ-product" and Colombeau's algebras, we first have to dismiss the positivity requirement for the δ-sequences employed to define $[U]\,[V]$. To this end we introduce the following further condition on a δ-sequence $\{\rho_j, j \in \mathbb{N}\} \subset \mathscr{D}(\mathbb{R}^n)$:

There is $M > 0$ so that for all $j \in \mathbb{N}$,

($\delta 4$) $$\int |\rho_j(x)|\, dx \leqq M.$$

Proposition 3. 3. *Suppose $U, V \in \mathscr{D}'(\mathbb{R}^n)$ satisfy condition* (MHO). *Then for any sequences of test functions $\{\rho_j, j \in \mathbb{N}\}$ and $\{\sigma_j, j \in \mathbb{N}\}$ satisfying* ($\delta 1$), ($\delta 2$), ($\delta 4$), *the limit*

$$\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V)$$

exists and equals $[U]\,[V]$.

Proof. Corresponding to the sequence $\{\rho_j, j \in \mathbb{N}\}$ there are ε_j and M so that ($\delta 1$), ($\delta 2$), ($\delta 4$) hold. It is possible to find functions $\chi_j \in \mathscr{D}(\mathbb{R}^n)$ so that $\chi_j \geqq 0$, $\rho_j + \chi_j \geqq 0$, $\chi_j(x) = 0$ for $|x| \geqq 2\varepsilon_j$, and for every $j \in \mathbb{N}$,

(4) $$\int \chi_j(x)\, dx = 2M, \quad \int (\rho_j(x) + \chi_j(x))\, dx = 1 + 2M.$$

Indeed, by mollifying $|\rho_j|$, for every $\delta > 0$ a function $\alpha_j \in \mathscr{D}(\mathbb{R}^n)$ may be constructed with $\alpha_j \geqq 0$, $\alpha_j(x) = 0$ for $|x| \geqq 2\varepsilon_j$, and

$$|\rho_j| \leqq \alpha_j + \delta, \quad \int \alpha_j(x)\, dx \leqq \int |\rho_j(x)|\, dx + \delta.$$

Taking functions $\beta_j \in \mathscr{D}(\mathbb{R}^n)$ with $\beta_j \geqq 0$, $\beta_j(x) = 1$ for $|x| \leqq \varepsilon_j$, $\beta_j(x) = 0$ for $|x| \geqq 2\varepsilon_j$, whose integrals are bounded by some constant $c > 0$ independent of j, we have that

$$|\rho_j| \leqq \alpha_j + \delta\beta_j, \quad \int (\alpha_j(x) + \delta\beta_j(x))\, dx \leqq M + \delta + c\delta.$$

Choosing δ small enough and possibly enlarging β_j we may achieve that

$$\int (\alpha_j(x) + \delta\beta_j(x))\, dx = 2M.$$

Thus $\chi_j = \alpha_j + \delta\beta_j$ has the desired properties.

Writing

$$\rho_j = (\rho_j + \chi_j) - \chi_j$$

we have decomposed ρ_j into the difference of two nonnegative functions whose supports tend to zero and whose integrals have the constant values given by (4). A similar construction can be done for $\{\sigma_j, j \in \mathbb{N}\}$. The proof is completed by observing that if $\{\varphi_j, j \in \mathbb{N}\}$, $\{\psi_j, j \in \mathbb{N}\}$ are sequences of test functions satisfying ($\delta 1$), ($\delta 3$) and

$$\int \varphi_j(x)\, dx = c_1, \quad \int \psi_j(x)\, dx = c_2 \quad \text{for all} \quad j \in \mathbb{N},$$

and if U, V satisfy (MHO), then $\lim_{j \to \infty} (\varphi_j * U)(\psi_j * V)$ exists and equals $c_1 c_2 [U][V]$. □

In Kamiński's notation, a sequence $\{\rho_j, j \in \mathbb{N}\}$ satisfying ($\delta 1$), ($\delta 2$) and

($\delta 5$) for every $p \in \mathbb{N}_0^n$ there is $M_p > 0$ so that
$\varepsilon_j^{|p|} \int |\partial^p \rho_j(x)|\, dx < M_p$ for all $j \in \mathbb{N}$

is said to *belong to the class* Δ. If $\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V)$ exists for all sequences ρ_j, σ_j belonging to the class Δ, this limit is called the *Δ-product of U and V*, see [5]. Since ($\delta 5$) implies ($\delta 4$), Proposition 3. 3 says that if the Mikusiński-Hirata-Ogata product $[U][V]$ exists, then Kamiński's Δ-product of U and V exists and equals $[U][V]$. The converse is not true. In fact, the Δ-product of $U = \delta(x)$ and $V = \sum_{m=1}^{\infty} \frac{1}{m^r} \delta\left(x - \frac{1}{m}\right)$ on $\mathbb{R}^1$ exists for $r > 2$ and equals 0, but $[U][V]$ does not exist (see Appendix).

Finally, if $\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V)$ exists for all sequences ρ_j, σ_j satisfying ($\delta 1$), ($\delta 2$), ($\delta 5$), then this limit also exists if we take in particular

$$\rho_j(x) = \sigma_j(x) = \varepsilon_j^{-n} \varphi\left(-\frac{x}{\varepsilon_j}\right)$$

for any $\varphi \in \mathscr{D}(\mathbb{R}^n)$ with $\int \varphi(x)\, dx = 1$. According to Colombeau's definition (Definition 3. 5. 2 of [2]) this means that the element $U \odot V$ in Colombeau's algebra $\mathscr{G}(\mathbb{R}^n)$ admits an *associated distribution* $U \tilde{\odot} V$ which equals

$$\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V).$$

Thus Colombeau's product $U \tilde{\odot} V$ is consistent with Kamiński's Δ-product and, by Proposition 3. 3, with the Mikusiński-Hirata-Ogata product (a case not covered by Theorem 3. 5. 7 of [2]). This in turn implies consistency with Ambrose' product $U \cdot V$. For the connection of the Mikusiński-Hirata-Ogata product with a product employing "restricted δ-sequences" (a subclass of the sequences satisfying ($\delta 1$), ($\delta 2$), ($\delta 3$)) we refer to Shiraishi [10].

Appendix

A counterexample. *Let* $U, V \in \mathscr{D}'(\mathbb{R}^1)$ *be given by*

$$U = \delta(x), \quad V = \sum_{m=1}^{\infty} \frac{1}{m^r}\, \delta\left(x - \frac{1}{m}\right) \quad \text{with} \quad r > 2.$$

Then $\lim_{j \to \infty} (\rho_j * U)(\sigma_j * V)$ *is equal to zero for all sequences* $\{\rho_j, j \in \mathbb{N}\}$, $\{\sigma_j, j \in \mathbb{N}\}$ *of test functions satisfying* (δ 1) (δ 2) (δ 5); *in particular, Kamiński's* Δ*-product of* U *and* V *exists. But the Mikusiński-Hirata-Ogata product* $[U][V]$ *does not exist.*

Proof. Let $\chi \in \mathscr{D}(\mathbb{R}^n)$. Then

$$\langle (\rho_j * U)(\sigma_j * V), \chi \rangle = \sum_{m=1}^{\infty} \frac{1}{m^r} \int \rho_j(x)\, \sigma_j\left(x - \frac{1}{m}\right) \chi(x)\, dx.$$

Let M_p, ε_j and M_p', ε_j' ($p \in \mathbb{N}_0$) be the positive numbers corresponding to ρ_j and σ_j, respectively, as required by (δ 1) and (δ 5). By (δ 1), the intersection of $\operatorname{supp} \rho_j$ and $\operatorname{supp} \sigma_j\left(. - \frac{1}{m}\right)$ is empty if $\frac{1}{m} \geqq \varepsilon_j + \varepsilon_j'$, that is, if $m \leqq l(j)$ where $l(j)$ denotes the largest integer in $(\varepsilon_j + \varepsilon_j')^{-1}$. On the other hand, ($\delta$ 5) implies that

$$\int |\rho_j(x)\, \sigma_j\left(x - \frac{1}{m}\right) \chi(x)|\, dx$$

is less or equal to both

$$\varepsilon_j^{-1} M_1 \cdot M_0' \sup_{x \in R} |\chi(x)| \quad \text{and} \quad M_0 \cdot (\varepsilon_j')^{-1} M_1' \sup_{x \in R} |\chi(x)|.$$

Thus, for some $c > 0$,

$$|\langle (\rho_j * U)(\sigma_j * V), \chi \rangle| \leqq c(\max\{\varepsilon_j, \varepsilon_j'\})^{-1} \sum_{m=l(j)+1}^{\infty} m^{-r}$$

$$\leqq c(\max\{\varepsilon_j, \varepsilon_j'\})^{-1} \frac{1}{r-1}\, l(j)^{1-r}$$

which tends to zero as $j \to \infty$, provided $r > 2$.

To prove that $[U][V]$ does not exist, it suffices to exhibit a sequence $\{\rho_j, j \in \mathbb{N}\}$ satisfying (δ 1), (δ 2), (δ 3), so that $\lim_{j \to \infty} U(\rho_j * V)$ does not exist. But

$$U(\rho_j * V) = \sum_{m=1}^{\infty} \frac{1}{m^r}\, \rho_j\left(-\frac{1}{m}\right) \delta.$$

Conditions (δ 1), (δ 2), (δ 3) cannot prevent us from taking ρ_j so that there is $m(j) > \varepsilon_j^{-1}$ with ε_j as in (δ 1) and $\rho_j\left(-\frac{1}{m(j)}\right) \geqq j \cdot m(j)^r$. Then

$$\sum_{m=1}^{\infty} \frac{1}{m^r}\, \rho_j\left(-\frac{1}{m}\right) \geqq j \to \infty. \qquad \square$$

Acknowledgements. The author wishes to thank N. Ortner and P. Wagner for helpful discussions and the referee for suggesting the present simpler and more general version of condition (δ 4).

Added in proof. The author recently learned that R. Wawak (Bull. Pol. Acad. Sci. **32** (1984), 179—183) has also proved the implication (A 1) ⇒ (A 2) and (A 3), using different arguments.

References

[1] *W. Ambrose*, Products of distributions with values in distributions, J. reine angew. Math. **315** (1980), 73—91.

[2] *J. F. Colombeau*, New generalized functions and multiplication of distributions, Amsterdam-New York-Oxford 1984.

[3] *Y. Hirata, H. Ogata*, On the exchange formula for distributions, J. Sci. Hiroshima Univ., A, **22** (1958), 147—152.

[4] *J. Horváth*, Topological vector spaces and distributions. I, Reading 1966.

[5] *A. Kamiński*, Convolution, product and Fourier transform of distributions, Studia Math. **74** (1982), 83—96.

[6] *J. Mikusiński*, Criteria of the existence and of the associativity of the product of distributions, Studia Math. **21** (1962), 253—259.

[7] *N. Ortner, P. Wagner*, Sur quelques propriétés des espaces $\mathscr{D}'_{L^p}$ de Laurent Schwartz, Boll. U.M.I. (6) **2-B** (1983), 353—375.

[8] *M. Reed, B. Simon*, Methods of modern mathematical physics. II: Fourier analysis, self-adjointness, New York-San Francisco-London 1975.

[9] *R. Shiraishi*, On the definition of convolutions for distributions, J. Sci. Hiroshima Univ., A, **23** (1959), 19—32.

[10] *R. Shiraishi*, On the value of a distribution at a point and multiplicative products, J. Sci. Hiroshima Univ., A-I, **31** (1967), 89—104.

[11] *R. Shiraishi, M. Itano*, On the multiplicative product of distributions, J. Sci. Hiroshima Univ., A-I, **28** (1964), 223—235.

[12] *M. E. Taylor*, Pseudodifferential operators, Princeton 1981.

[13] *J. Tysk*, On the multiplication of distributions, Uppsala Univ. Math. Dept., Project Report 1981.

[14] *V. S. Vladimirov*, Generalized functions in mathematical physics, Moscow 1979.

Institut für Mathematik und Geometrie, Universität Innsbruck, Technikerstraße 13, A-6020 Innsbruck

Eingegangen 27. September 1984

Valuations of function fields from the geometrical point of view

By *Ludwig Bröcker* at Münster and *Heinz-Werner Schülting* at Dortmund

Introduction

Let C be a smooth integral complete algebraic curve, defined over a field k. Then C is determined by its function field F/k. In particular the underlying space corresponds to the space $\mathrm{val}(F/k)$ of all k-valuations of F via the map $C \to \mathrm{val}(F/k)$; $x \to \mathcal{O}_{(x,C)}$[1]) and each $v \in \mathrm{val}(F/k)$ is discrete of rank 1 with a residue field F_v such that F_v/k is finite.

Now let X be an algebraic integral k-variety of dimension $\geqq 2$. There is no longer an obvious connection between $\mathrm{val}(F/k)$, $F = k(X)$, and the geometry of X. Moreover a valuation v of F/k may be

- of arbitrary $\mathbb{Q}$-rank between 1 and $\dim X$,
- not discrete,
- with a residue field F_v which is not finitely generated over k,
- such that $\mathrm{rank}_{\mathbb{Q}}(v) + \dim(v) < \dim X$

and so on, in other words, v may be arbitrarily complicated. Therefore valuation theory seems to be less useful for the study of algebraic varieties of higher dimensions. However, the valuations v of F/k should play a role in the birational geometry of F/k, that is, the valuation v should be found somehow in the category of all models X of F/k. This is what we shall investigate.

On the way we want to understand the above phenomena and to gain a perspective on the entire space $\mathrm{val}(F/k)$ from this geometrical point of view.

There are two sources for the new interest in Krull valuations (valuations of arbitrary rank):

1) Recent investigations, where algebraic geometry and commutative algebra is amalgamated with a further structure: Partially ordered commutative algebra [Bf], real algebraic geometry, the real spectrum [Co-CR] and formally p-adic fields [P-R]...

[1]) Compare our "notations and conventions"-table at the end of this introduction.

2) General local-global principles in the theory of forms: Representation of elements by n^{th} powers [B], weak isotropy for quadratic forms [Br1], and for forms of higher degree [B], the "triumvirat" orderings, valuations and quadratic forms [L].

We hope to present a general frame work, under which the valuations occurring above can be treated in a geometrical situation. See § 4 and [S2] for successful applications.

Note, that the valuations, which are involved in 1) and 2) are such that their residue fields belong to a special class, namely the formally real or formally p-adic fields respectively.

In § 1 we shall define a proper class T of fields by axioms in terms of model theory. However, our use of that branch of mathematics is very modest and restricted to § 1. In all the work we study T-valuations v of function fields F/k, i.e. v is trivial on k, $k \in T$ and the residue field F_v belongs to the class T as well.

All fields in T will be of char 0.

In § 2 we characterize T-central points, that is centres of T-valuations of F/k on models X of F/k. We construct T-valuations with prescribed centres, finitely generated value groups, dimensions of the residue field and to some extend with prescribed residue fields. We do this also for chains of centres and chains of valuations. Our main result is Theorem (2. 12) together with (2. 10) which generalizes recent work of Andradas [A].

In § 3 we study the ring $H = H(F/k, T) \subset F$. This is the intersection of all T-valuation rings of F/k. It is well known, that $H = k$ if k is algebraically closed and T is the class of all fields. But if k is existentially closed in T but not algebraically closed, the ring H is very interesting: It is a Prüfer domain. It can be approximated by rings of functions, which are regular on all k-rational points for models X of F/k. The class group of H can be approximated in a similar way by class groups of so called T-divisors on models X of F/k. We describe the Zariski space $\mathrm{val}(F/k, T)$ of all T-valuations of F/k by an algebraic limit and give some examples.

In the short concluding § 4 we present an application. For instance this makes Becker's criterion for the representation of field elements $f \in F$ by sums of n^{th} powers in F more computational in the case that F is a function field over $\mathbb{R}$ (compare [B] and [S2]).

We shall freely use the language of the modern algebraic geometry and the material of Chapter VI in "Zariski-Samuel" [Z-S].

However, our main tool and permanent reference is Hironakas work on the resolution of singularities [H]. In particular we use Main Theorem I ([H], Chapter 0, § 3), Main Theorem II and its Corollary 1 (elimination of points of indeterminacy) ([H], Chapter 0, § 5). We hope that our way to understand valuations of function fields justifies the use of these powerful and deep results.[2] [3]

---

[2] Recently Kuhlmann and Prestel developed methods for the investigation of valuations of function fields without using Hironakas work, replacing it by the Ax-Kochen-Ershov principle [K-P]. With these methods one can also prove results like Theorem 2. 12 and Corollary 4. 3.

[3] We thank the referee, who read this manuscript carefully and gave us some hints. In particular the suggestion of footnote 4) is due to him.

Notations and conventions

1. k a commutative field of characteristic 0: k-variety V = separated integral scheme of finite type over k. $x(V)$ = generic point of V. Subvarieties and morphisms are always k-subvarieties and k-morphisms. $W < V$ means W is a closed subvariety of V. For $U \subset V$ we denote by $\bar{U}$ the Zariski closure of U in V. $\dim x = \dim \bar{x}$ for $x \in V$.

$\mathcal{O}_{(x,V)}$ = local ring at x. $\mathcal{M}_{(x,V)}$ = maximal ideal of $\mathcal{O}_{(x,V)}$.

V is called smooth, if $\mathcal{O}_{(x,V)}$ is regular for all $x \in V$ (this is defined in the same way, if V is not necessarily integral. Since $\operatorname{char} k = 0$ we need not distinguish between smooth and non-singular).

$V(K)$ = set of K-points for a field extension K/k.

2. F/K a field extension: $d(F/K)$ = degree of transcendency.

A valuation v or a place φ of F/K is always trivial on K.

B_v = valuation ring, m_v = maximal ideal, F_v = residue field, $|F|_v$ = value group, $\dim v = d(F_v/K)$.

For places φ we use the corresponding notations. They are always surjective.

3. K a commutative field: $K(t)$ = purely transcendental extension of K with $d(K(t)/K) = 1$,

$K((t))$ = field of formal power series over K,

$K\{t\}$ = field of generalized power series over K ([Z-S], p. 101),

$\bar{K}$ = algebraic closure of K.

4. U an abelian group: $\operatorname{rank}_{\mathbb{Q}} U = \dim_{\mathbb{Q}} \mathbb{Q} \underset{\mathbb{Z}}{\otimes} U$.

§ 1. General assumption

Let $\mathcal{T}_0$ be the first order theory of fields with respect to the language $\mathcal{L}_0 = \{+, -, \cdot, (\,)^{-1}, 0, 1\}$. We assume, that there is given an expanded language $\mathcal{L} \supset \mathcal{L}_0$ and a set of axioms in $\mathcal{L}$ which contains the field axioms. Thus we get an extended theory $\mathcal{L}, \mathcal{T}$. By T we denote the class of all models of $\mathcal{T}$. So T is a class of fields with some supplementary structure. We assume, that the following conditions hold for the class T.

A 0. *Each field $F \in T$ has characteristic* 0.

A 1. *If $F \in T$ and $k \subset K \subset F$ are substructures with $k \in T$, then $K \in T$.*

From these one gets by the Löwenheim-Skolem-Tarski theorem [C-K], Theorem 3.1.5.

A 2'. *If $K \in T$ then $K(t) \in T$ too (for a suitable interpretation of $\mathcal{L} \setminus \mathcal{L}_0$ on $K(t)$).*

If not otherwise mentioned we need the stronger property

A 2. *If $K \in T$ then $K((t)) \in T$ too (for a suitable interpretation of $\mathcal{L} \setminus \mathcal{L}_0$ on $K((t))$).*

(1. 1) Remark. A2 is not a consequence of A0 and A1. For instance consider the plane curve C defined by $f = X_1^3 + X_2^3 - 1 = 0$ with the only $\mathbb{Q}$-rational points $p_1 = (1, 0)$ and $p_2 = (0, 1)$ which are simple. They are on the line given by $g = X_1 + X_2 - 1 = 0$. Now adjoin to the axioms of fields of char 0 the further axiom:

$\forall x, y: f(x, y) = 0 \rightarrow g(x, y) = 0$. Then $\mathbb{Q} \in T$ and A0, A1 hold for T but A2 fails to hold:

$\mathbb{Q}((t))$ is an extension of $\mathbb{Q}(C)$ since the valuation ring $\mathcal{O}_{(p_1, C)}$ of $\mathbb{Q}(C)$ can be imbedded into $\mathbb{Q}((t))$, but it is clear that $\mathbb{Q}(C) \notin T$.

For all this work let k be a fixed field in our class T. Sometimes we assume that k is existentially closed in T, as in the following examples.

(1. 2) Examples. a) $\mathcal{T}$ is the theory of the fields of char 0, hence T is the class of all fields of char 0, and k is an algebraically closed field of T.

b) $\mathcal{T}$ is the theory of the formally real fields and k is a fixed real closed field of T.

c) For $p, d \in \mathbb{N}$, p a prime, let $\mathcal{T}$ be the theory of the formally p-adic fields of rank d and $k \in T$ a fixed p-adically closed field of rank d.

d) $\mathcal{T}$ is the theory of the ordered fields and $k \in T$ is a fixed maximal ordered (thus real closed) field.

In these examples the existentially closed fields are just the maximal algebraic fields in the class T. This follows for a) by Hilbert's Nullstellensatz, for b) and d) by the Artin-Lang theorem [Ar] and for c) by results of Kochen and Roquette [P-R], Theorem 7. 8.

In the only statement of this section, which is nearly tautological, we derive algebraic properties from the notion of "existentially closed" [B-J], Theorem (1. 1).

(1. 3) Proposition. *For $\mathcal{T}$ and T as above and $k \in T$ consider the following properties*:

a) *k is existentially closed with respect to $\mathcal{T}$.*

b) *If A is a finitely generated commutative integral k-algebra such that its field of fractions $F \in T$, then A admits a k-algebra homomorphism $\varphi: A \rightarrow k$.*

c) *If V is an affine k-variety such that $k(V) \in T$, then the set of its non-singular k-rational points is Zariski-dense in V.*

Then a) $\rightarrow$ b) $\leftrightarrow$ c). *Moreover, if $\mathcal{T}$ is the theory of all fields, then* b) $\rightarrow$ a) too.

§ 2. Valuations with given data

Assume that $\mathcal{T}$, T and k are given as in § 1. However, so far the field k need not to be existentially closed.

(2. 1) Proposition. *Let V be a k-variety and $W \subset V$ a subvariety such that $x(W)$ is a simple point in V. Let $K/k(W)$ be a function field with*

$$d(K/k(W)) < d(k(V)/k) - d(k(W)/k) - 1.$$

Then there exists a place φ of $k(V)/k$ with residue field $k(V)_\varphi \cong K$ and centre $x(W)$.

Proof. We may assume that V is affine and V, W both are smooth. Then the blowing up $\pi\colon \tilde{V} \to V$ of V in W is smooth too and its exceptional divisor D is $W \times \mathbb{P}^r$, $r = \dim V - \dim W - 1$. We have $U = \operatorname{spec}(k[W][X_1, \ldots, X_r]) \subset D$, U open. By assumption $r > d(K/k(W))$, hence K is of the form $K = k(W)(x_1, \ldots, x_r)$.

For the kernel I of the canonical map $k[W][X_1, \ldots, X_r] \to k(W)(x_1, \ldots, x_r)$ one has $I \cap k[W] = \{0\}$. Hence I defines a subvariety Z of D with $k(Z) \cong K$ and $\pi(x(Z)) = x(W)$. Now it is sufficient to find a place of $k(\tilde{V})/k$ with centre Z in $\tilde{V}$ and residue field $k(\tilde{V})_\varphi \cong k(Z) \cong K$. That means, we look for a valuation ring B of $k(\tilde{V})$ which dominates $\mathcal{O}_{(x,\tilde{V})}$, $x = x(Z)$, such that the map $\mathcal{O}_{(x,\tilde{V})} \to B$ defines an isomorphism on the residue fields. Since $\mathcal{O}_{(x,\tilde{V})}$ is regular there exists even a discrete valuation ring B of rank $l = \dim \tilde{V} - \dim Z$ with that property as it is well-known.

Let us introduce the main notions of this section.

(2. 2) Definitions. Let V be a k-variety. A point $x \in V$ is called T-point, if $k(x) \in T$. The set of all T-points of V will be denoted by $V(T)$. We call V a T-variety, if $x(V)$ is a T-point.

For $K \in T$ a place (or valuation) φ of K is named T-place, if $K_\varphi \in T$.

A point $x \in V$ is called T-central or central, if x is centre of a T-place of $k(V)$. We write $V\{k\}$ for the set of all k-rational central points.

Note that by A1 a central point is always a T-point. Moreover, by (2. 1) one has $V_{\text{simpl}}(k) \subseteq V\{k\} \subseteq V(k)$.

(2. 3) Proposition. *Assume, that the k-variety V admits a simple T-point x. Then V itself is a T-variety.*

Proof. Let b be a regular parameter of $O := \mathcal{O}_{(x,V)}$. Then $K := O_{(b)}/bO_{(b)} \in T$ by induction on the dimension of V. On the other hand $O_{(b)}$ is a discrete valuationring of $k(V)$, hence $k(V)$ imbeds into $K((t))$. Using A2 we get $k(V) \in T$.

One gets another proof of (2. 3) by blowing up V at x. A third argument:

$$
\begin{aligned}
&K = k(x),\\
&x \in T,\\
&x \text{ simple} \Rightarrow \mathcal{O}_{(x,V)} \hookrightarrow K((t_1, \ldots, t_r)) \subseteq K((t_1))(\ldots)((t_r)) \in T\\
&\qquad\qquad \Rightarrow \operatorname{Quot}(\mathcal{O}_{(x,V)}) \in T.
\end{aligned}
$$

Remark (1. 1) shows that the use of A2 is essential.

(2. 4) Corollary. *Let F/k be a function field which admits a T-place. Then F is contained in T.*

Proof. Choose a smooth projective model X of F and let φ be a T-place. The centre of φ on X is a regular T-point, hence X is a T-variety and $F \in T$.

Here for the first time we used the desingularization [H].

Recalling our general assumption and the notation of (2. 2) we characterize the central points of a variety as follows.

(2. 5) Theorem (Characterization of central points). *Let k be existentially closed in T and let X be a projective k-variety, $F := k(X)$ and $z \in X$ with Zariski-closure $\bar{z} = Z$.*

The following statements are equivalent:

a) *The point z is central.*

b) *There exists a surjective birational morphism $f: Y \to X$, Y a projective k-variety, and a simple T-point $y \in Y$ with $f(y) = z$.*

c) *$Z \cap X\{k\}$ is dense in Z.*

Proof. a) $\Rightarrow$ c). By assumption we have a T-place φ of F with centre z. Let $f: Y \to X$ be a desingularization of X and y the centre of φ on Y, in particular $f(y) = z$. The closure $V = \bar{y}$ of y in Y is a T-variety, hence by (1. 3) $V(k)$ is dense in V. Using (2. 1) we get $Y(k) = Y\{k\}$, thus $Y\{k\} \cap V = V(k)$ is dense in V and c) follows.

c) $\Rightarrow$ b). Let $f: Y \to X$ be a projective desingularization of X and let V be the Zariski-closure of $Y\{k\} \cap f^{-1}(Z)$. Then $f(V) \subset Z$ and since the fibres of the points of $X\{k\} \cap Z$ admit central points (in fact $f(\varphi(Y)) = \varphi(X)$ if $\varphi(Y)$ is the centre of a T-place φ on Y and $\varphi(X)$ is its centre on X) we get $X\{k\} \cap Z \subset f(V)$. Thus, by assumption $f(V)$ is dense in Z. Then for a suitable component V_1 of V we also have: $f: V_1 \to Z$ is dominant and $Y\{k\} \cap V_1$ is dense in V_1. In particular, V_1 contains regular rational points, hence V_1 is a T-variety, $x(V_1)$ is a T-point, and $f(x(V_1)) = z$.

b) $\Rightarrow$ a). This follows from (2. 1).

After we have characterized central points it might be interesting to consider chains of central points in the following sense.

(2. 6) Definition. Let X be a k-variety. A chain of points

$$x(X) = x_{s+1},\ x_s,\ \ldots,\ x_1 \in X$$

such that x_{i+1} generalizes x_i properly for $i = 1, \ldots, s$ is called a T-central chain, if there are T-valuations v_i of $F = k(X)$, v_{i+1} coarser than v_i, v_{s+1} trivial such that v_i has centre x_i for $i = 1, \ldots, s$.

We are going to characterize T-central chains and to construct chains of places for a given T-central chain and further given data such as value groups, dimensions of residue fields and even to a certain extent for given residue fields. We need the following

(2. 7) Lemma. *Let $f: V \to V'$ be a dominant k-morphism of affine k-varieties. Let $x \in V$ be a simple point with $f(x) = x'$ and $\dim x = \dim x' < \dim V'$. Then there exists a generalization y of x such that y is simple in V, x simple in $\bar{y}$ and*

$$\dim y = \dim f(y) = 1 + \dim x.$$

Proof. This is true if $\dim V = \dim x + 1$. Hence by induction on n we must only find a subvariety H of dimension $n-1$, $n = \dim V$ such that x is a simple point of H, $x(H)$ a simple point of V and $f(H)$ is not contained in the Zariski closure W' of x'. Let $W_1, \ldots, W_r$ be the components of $W = f^{-1}(W')$ with $x \in W_i$ and $\dim W_i = n-1$ (if they exist). We may assume, that $\dim x < n-1$, hence $\dim \mathcal{O}_{(x,W_i)} > 0$. Therefore we have an element $u \in \mathcal{M}_{(x,V)}$ such that $u + \mathcal{M}^2_{(x,V)}$ is not in the kernel of the canonical surjection $\mathcal{M}_{(x,V)}/\mathcal{M}^2_{(x,V)} \to \mathcal{M}_{(x,W_i)}/\mathcal{M}^2_{(x,W_i)}$ for $i = 1, \ldots, r$. Now choose for H the component of the variety of u such that $x \in H$.

(2. 8) Theorem (Characterization of T-central chains). *Let k be existentially closed in T and let X be a k-variety with a chain of points $x(X) = x_{s+1}, x_s, \ldots, x_1$ in X.*

The following are equivalent:

a) *$x_{s+1}, x_s, \ldots, x_1$ is a T-central chain.*

b) *There is a birational k-morphism $f: Y \to X$ of k-varieties and a chain $x(Y) = y_{s+1}, y_s, \ldots, y_1$ of points in Y, $y_i \in \bar{y}_{i+1}$ such that $x_i = f(y_i)$, $\dim y_i = \dim x_i$, y_1 is a T-point and y_i is a simple point of $\overline{y_{i+1}}$ for $i = 1, \ldots, s$.*

Proof. a) $\to$ b). Let $g: Z \to X$ be a desingularization. Assume that we have already smooth closed subvarieties $Z = Z_{s+1} > Z_s > \cdots > Z_r$ with generic points $z_{s+1}, \ldots, z_r$ and a T-central chain $z_{s+1}, z_{r-1}, \ldots, z_1$ such that $z_{r-1} \in Z_r$ and $g(z_i) = x_i$ for $i = 1, \ldots, s+1$. According to Hironaka's Main Theorem I (loc. cit.) there is a morphism $g': Z'_{r-1} \to Z_{r-1} := \overline{z_{r-1}}$ which is a composition of monoidal transformations in smooth centres such that Z'_{r-1} is smooth. Apply the same monoidal transformations to Z_j, $r \leqq j \leqq s+1$. Since this process preserves the closed immersions $Z_j < Z_{j+1}$ ([Ha], II, Corollary 7.15) we obtain a chain $Z'_{s+1} > \cdots > Z'_r > Z'_{r-1}$ of smooth varieties and a birational morphism $g': Z'_{s+1} \to Z_{s+1}$. Set $z'_i := x(Z'_i)$ for $s+1 \geqq i \geqq r-1$. The T-central chain $z_{s+1}, z_{r-2}, \ldots, z_1$ can be lifted to the T-central chain $z'_{s+1}, z'_{r-2}, \ldots, z'_1$ of Z' with $z'_{r-2} \in Z_{r-1}$ (consider the centres of the valuations, which define the original T-chain). Now we get $g' \circ g(z'_i) = x_i$ for all $i = 1, \ldots, s+1$. Repeating this procedure we obtain a chain $W_{s+1} > W_s > \cdots > W_1$ of smooth k-varieties and a birational morphism $h: W_{s+1} \to X_{s+1}$ such that $h(w_i) = x_i$ for $w_i = x(W_i)$, $i = 1, \ldots, s+1$. By construction it is clear, that each W_i is a T-variety. We have not yet $\dim w_i = \dim x_i$. In order to get this we choose a closed T-point $y_0 \in W_1$ which exists by (1. 3). Then by repeated application of (2. 7) we get a chain of subvarieties $Y_i < W_i$, $Y = Y_{s+1} > Y_s > \cdots > Y_1 > y_0$ such that for $y_i = x(Y_i)$ and $f = h|Y_{s+1}$ one has $h(y_i) = x_i$, y_{i-1} is a simple point of Y_i and $\dim y_i = \dim x_i$ for $i = 1, \ldots, s+1$. Since y_0 is a T-point, by (2. 3) the same holds for y_1 and the other y_i which proves the assertion.

b) $\to$ a). This follows easily from (2. 1).

In the following two propositions let $\mathcal{T}$ be the theory of the formally real fields, so k is real closed. Here one has a topological characterization of T-central points and chains. To explain this we assume, that the reader is aquainted a little with semialgebraic geometry (see [D-K], [Co-CR]). For a semialgebraic set $S \subset X(k)$, X a k-variety, we denote by S^c the set of all points $x \in S$ such that the local dimension of S at x equals $\dim S$. If S is Zariski-dense in X, then S^c consists of the limit points of those interior points of S (with respect to the strong topology of $X(k)$) which are simple in X. In particular S^c is semialgebraic and closed with respect to the strong topology of $X(k)$.

(2. 9) Proposition. *If X is real, then $X\{k\} = X(k)^c$.*

Proof. [E].

By (1. 3) X is real iff $\overline{X(k)} = X$. Then $\overline{X(k)^c} = X$ too. Now let

$$x(X) = x_{s+1}, x_s, \ldots, x_1,$$

a chain of points in X, $\overline{x_i} = X_i$. We define a semialgebraic set $S_i \subset X_i(k)$ inductively as follows: $S_{s+1} := X_{s+1}(k)^c = X(k)^c$ and $S_i = (S_{i+1} \cap X_i)^c$ for $i = 1, \ldots, s$.

(2. 10) Proposition. *Let $\mathscr{T}$ be the theory of the formally real fields. $x_{s+1}, x_s, \ldots, x_1$ a chain of points in X and S_i, X_i as above. Then the following properties are equivalent.*

a) *$x_{s+1}, x_s, \ldots, x_1$ is a T-central chain.*

b) *S_i is Zariski-dense in X_i for $i = 1, \ldots, s+1$.*

Proof. a) $\rightarrow$ b). Let v_i, $i = 1, \ldots, s$ be real valuations of $F = k(X)$ with centre x_i such that v_{i+1} is coarser than v_i. Choose an ordering P of F which is compatible with v_1 and therefore also compatible with the other v_i. Then as an element of the real spectrum of X the point P has a chain of specializations $P = P_{s+1}, P_s, \ldots, P_1$ where the centre of P_i is x_i ([Br], Proposition 2. 13). Then the corresponding ultrafilter $f(P)$ on $X(k)$ converges to the chain $x_{s+1}, x_s, \ldots, x_1$ (loc. cit. 4. 1—4. 4) which shows b).

b) $\rightarrow$ a). It is not hard to construct an ultrafilter U on $X(k)$ which converges to the chain $x_{s+1}, x_s, \ldots, x_1$. The corresponding element $P = p(U)$ in the real spectrum of X has a chain of specializations $P = P_{s+1}, P_s, \ldots, P_1$ where again the centre of P_i is x_i. By Brumfiels place extending theorem ([Bf], Proposition 7. 7. 10) we find a chain of valuations $v_s, \ldots, v_1$ of F, v_1 compatible with P and v_{i+1} coarser than v_i such that $x_i =$ centre of v_i.

Let us return to an arbitrary theory $\mathscr{T}$ with the class of models T as in § 1. We assume, that k is existentially closed in T and consider a k-variety X with a T-central chain $x(X) = x_{s+1}, x_s, \ldots, x_1$, $n_i = \dim x_i$, $n_{s+1} = n = \dim X$. We look for a corresponding sequence of function fields $F_i/k(x_i)$ and places φ_i: $F_{i+1} \rightarrow F_i \cup \infty$ with $x_i =$ centre of $\psi_i = \varphi_i \circ \varphi_{i+1} \circ \cdots \circ \varphi_s$, with prescribed values for $d_i := d(F_i/k)$ and prescribed value groups U_{i+1}/U_i of φ_i where U is a totally ordered abelian group, $\{1\} = U_1 \subset \cdots \subset U_s \subset U_{s+1} = U$ a sequence of convex subgroups. Of course, there are some restrictions. First of all (if our places are assumed to be non-trivial) one has $n = d_{s+1} > d_s > \cdots > d_1$ and $d_i \geqq n_i$. Secondly it is clear, that $\operatorname{rank}_Q U_{i+1}/U_i$ must be $\leqq d_{i+1} - d_i$. Note, that by (2. 8) it makes no restriction to assume, that the given T-central chain is regular, that means, x_i is a simple point of X_{i+1} for $i = 1, \ldots, s$.

One might ask, whether one can also prescribe the residue fields F_i. We shall do a little in that direction. However, there are restrictions too, as we see by the following

(2. 11) Proposition. *Let v be a T-valuation of $F = k(X)$ of codimension 1 with centre x, x a simple point of X. Then $|F|_v \cong \mathbb{Z}$ and for the residue field F_v one has*

a) *If the codimension of $x = 1$, then $F_v \cong k(x)$.*

b) *If the codimension of $x \geqq 2$, then $F_v \cong F_1(t)$ for a suitable function field $F_1/k(x)$.*

Proof. In the case a) the valuation ring of v is $\mathcal{O}_{(x,X)}$ which yields the assertion. In case b) we may assume that X is smooth (otherwise replace it by a smooth affine neighbourhood of x). Starting from X we obtain by finitely iterated blowing up along smooth centres ([H], p. 144) a model Y such that for the centre y of v in Y and its residue field $k(y)$ we have $F_v \supset k(y) \supset k(x)$ and $d(k(y)/k) = d(F_v/k) > d(k(x)/k)$ (compare Proposition (3. 9)). So assume that we get Y by blowing up Z at the centre W and that for the centre z of v in Z we still have $\dim z < \dim Z - 1$. Now if $z \notin W$ then y is the strict transform of z so $\dim y = \dim z < \dim X - 1$. If z is a non-generic point in W then $\dim y < \dim X - 1$ too. In the remaining case $z \in W$, z generic, the fibre of z is a projective space over $k(z)$ which proves our assertion.

Now let us state our main existence theorem.

(2. 12) Theorem. *Let k be existentially closed, X a k-variety and $x(X) = x_{s+1}$, $x_s, \ldots, x_1$ a T-central chain such that x_i is a simple point of $X_{i+1} = \bar{x}_{i+1}$. Suppose, there is given:*

a) *A sequence of integers $n = d_{s+1} > d_s > \cdots > d_1$ and $d_i \geqq n_i := \dim x_i$ for $i = 1, \ldots, n$.*

b) *A finitely generated totally ordered abelian group U with convex subgroups $\{1\} = U_1 \subset \cdots \subset U_s \subset U_{s+1} = U$ such that $\operatorname{rank}_{\mathbb{Q}} U_{i+1}/U_i \leqq d_{i+1} - d_i$ for $i = 1, \ldots, n$.*

c) *For the case that $d_2 > d_1 + 1$ a function field $F_1/k(x_1)$ with $d(F_1/k) = d_1$.*

Then for $i = 1, \ldots, s$ there are function fields $F_i/k(x_i)$ with $d(F_i/k) = d_i$ (where F_1 is possibly prescribed by c) and $F_{s+1} = F = k(X)$) and places $\varphi_i \colon F_{i+1} \to F_i \cup \infty$ with value groups $\cong U_{i+1}/U_i$ such that $\psi_i := \varphi_i \circ \cdots \circ \varphi_s \colon F \to F_i \cup \infty$ has centre x_i in X.

Proof in five steps: 1) $s = 1$, $F_1 = k(x_1)$ and U is isomorphic to a subgroup of $(\mathbb{R}, +)$. We may assume, that X is affine and defined by polynomials $f_{n+1}, \ldots, f_m \in k[X_1, \ldots, X_m]$.[4]) Then x_1 corresponds to a prime ideal $P_1 \supset P = (f_{n+1}, \ldots, f_m)$ of $k[X_1, \ldots, X_m]$. Moreover for a suitable numeration of the coordinates we have

$$\det\left(\frac{\partial(f_{n+1}, \ldots, f_m)}{\partial(X_{n+1}, \ldots, X_m)}(u_1, \ldots, u_m)\right) \neq 0$$

where $u_1, \ldots, u_m$ are the coordinates of x_1, that is, $u_i = X_i + P_1$ in $k[X_1, \ldots, X_m]/P_1$. The formal implicit function theorem provides power series

$$\gamma_{n+1}, \ldots, \gamma_m \in k(x_1)\,[[t_1 - u_1, \ldots, t_n - u_n]]$$

with $u_j =$ constant term of γ_j and $f_j(t_1, \ldots, t_n, \gamma_{n+1}, \ldots, \gamma_m) = 0$ for $j = n+1, \ldots, m$. Recall, that $d = d_1 = n_1 = \dim x_1$. We may arrange the u_i such that $u_{n-d+1}, \ldots, u_n$ are algebraically independent over k. Now choose a $\mathbb{Z}$ basis $\alpha_1, \ldots, \alpha_r$ of $U \subset \mathbb{R}$ with $\alpha_1 > \alpha_2 > \cdots > \alpha_r > 0$ and power series $\delta_{r+1}, \ldots, \delta_n \in k(x_1)\,[[t^{\alpha_1}]]$ such that the generalized power series ([Z-S], p. 101)

$$\delta_1 = t^{\alpha_1} + u_1, \ldots, \delta_r = t^{\alpha_r} + u_r, \delta_{r+1}, \ldots, \delta_n$$

---

[4]) In fact, locally X can be described as a hypersurface $f_m = 0$, smooth at x_1. By this observation one can simplify the following computations a little.

are algebraically independent over $k(x_1)$ and $\operatorname{ord}(\delta_{r+j}-u_{r+j})=2\alpha_1$ for $1\leqq j\leqq n-r$. Here ord denotes the canonical valuation of the field $k(x_1)\{t\}$ of generalized power series. Such series δ_i exist according to MacLane and Shilling [M-S]. Since $\delta_j(0)=u_j$ for $1\leqq j\leqq n$ we may substitute the power series $\delta_1,\ldots,\delta_n$ into $\gamma_{n+1},\ldots,\gamma_m$ and obtain generalized power series

$$\delta_{n+i}(t):=\gamma_{n+i}(\delta_1(t),\ldots,\delta_n(t)).$$

Consider the homomorphism

$$\varphi\colon k[X_1,\ldots,X_m]\to k(x_1)[\delta_1,\ldots,\delta_m]$$

defined by $\varphi(X_i)=\delta_i$ for $i=1,\ldots,m$. We claim that $\ker\varphi=P=(f_{n+1},\ldots,f_m)$. By construction $P\subset\ker\varphi$. On the other hand the power series $\delta_1,\ldots,\delta_n$ are algebraically independent over $k(x_1)$, thus

$$k[X_1,\ldots,X_n]\cap\ker\varphi=\{0\}$$

and therefore

$$\dim(k[X_1,\ldots,X_m]/\ker\varphi)\geqq n=\dim(k[X_1,\ldots,X_m]/P),$$

hence $P=\ker\varphi$. Thus φ induces an imbedding $F=k(X)\to k(x_1)\{t\}$ which yields a rank 1 valuation v of F with centre x_1 and residue field $k(x_1)$. It remains to show that $|F|_v=U$. By construction $|F|_v\subset U$. Recall that $r+d\leqq n$ and that $u_{n-d+1},\ldots,u_n$ are algebraically independent over k. So for $1\leqq j\leqq r$ let be $h_j\in k[T_j,T_{n-d+1},\ldots,T_n]$ a polynomial of minimal degree in T_j with $h_j(u_j,u_{n-d+1},\ldots,u_n)=0$.

Then $\delta(t):=h_j(\delta_j,\delta_{n-d+1},\ldots,\delta_n)$ has no constant term and we note that $\frac{\partial}{\partial T_j}h_j(u_j,u_{n-d+1},\ldots,u_n)=:u\neq 0$. Now

$$\frac{d\delta}{dt}=\frac{d}{dt}h_j(\delta_j,\delta_{n-d+1},\ldots,\delta_n)=\sum_{i=n-d+1}^{n}\frac{\partial}{\partial T_i}h_j(\delta_j,\delta_{n-d+1},\ldots,\delta_n)\cdot\frac{d}{dt}\delta_i$$
$$+\frac{\partial}{\partial T_j}h_j(\delta_j,\delta_{n-d+1},\ldots,\delta_n)\cdot\frac{d}{dt}\delta_j.$$

Denoting the two terms on the right hand side by $\rho_1(t)$ and $\rho_2(t)$ we see that the leading term of $\rho_2(t)$ is ut^{α_j} whereas the exponent of the leading term of $\rho_1(t)\geqq 2\alpha_1>\alpha_j$ or $\rho_1(t)=0$. Therefore

$$v(h_j(\bar{X}_j,\bar{X}_{n-d+1},\ldots,\bar{X}_n))=\alpha_j$$

for $j=1,\ldots,r$ and $\bar{X}_i:=X_i+P$.

Now the canonical place φ_1 of v fullfills the assertion.

2) $s=1$, $F_1=k(x_1)$ but U arbitrary. Assume, that $1=H_1\subset H_2\subset\cdots\subset H_{l+1}=U$ are the convex subgroups of U. Then

$$\sum_{i=1}^{l}\operatorname{rank}_Q H_{i+1}/H_i=\operatorname{rank}_Q U\leqq n-d$$

where again $d=d_1=n_1=\dim x_1$. Therefore we find points

$$x_1=z_1,\ldots,z_{l+1}=x_2=x(X),$$

z_i a simple point of $\overline{z}_{i+1}$ such that $\operatorname{rank}_{\mathbb{Q}} H_{i+1}/H_i \leqq \dim z_{i+1}-\dim z_i$. Since H_{i+1}/H_i imbeds into $\mathbb{R}$, + by step 1) we have a place $\chi: F\to k(z_l)$ with value group U/H_l and by induction on l we have a place $\chi': k(z_l)\to k(z_1)=k(x_1)$ with valuegroup H_l. Then $\varphi_1:=\chi'\circ\chi$ is the place we look for.

3) $s=1$. Let $f: Y\to X$ be the blowing up of X at X_1 with exceptional divisor $E=f^{-1}(X_1)$. If $d_1=n-1$ set $y_1=x(E)$. If $d_1<n-1$ for a given function field $F_1/k(x_1)$ with $d(F_1/k)=d_1$ we find as in the proof of (2. 1) a point $y_1\in E$ with $k(y_1)\cong F_1$ and $f(y_1)=x_1$. Now we apply 2) to the T-central chain $x(Y)=y_2, y_1$.

4) Now we assume that $s>1$ and prove the general case by induction on s. First consider the case $d_s=n_s$. By step 1) we find a place $\varphi_s: F\to k(x_1)\cup\infty$ with valuegroup U/U_s. Now $x_s,\ldots,x_1$ is a regular central T-chain of X_s. By induction we find function fields $F_i/k(x_i)$ and places $\varphi_i: F_{i+1}\to F_i\cup\infty$ for $i=1,\ldots,s-1$ such that the theorem holds for X_s, the T-central chain $x_s,\ldots,x_1$, the sequence $d_s>\cdots>d_1$, the chain of groups $U_1\subset\cdots\subset U_s$ and eventually the given field F_1. Then the chain of places $\varphi_s, \varphi_{s-1},\ldots,\varphi_1$ does the job.

5) Reduction to the situation in step 4). Choosing eventually an affine neighbourhood of x_1 we may assume, that all X_i are smooth. Let $f: Y\to X$ be the blowing up of X at X_s. The exceptional divisor E is isomorphic to $X_s\times\mathbb{P}^{n-n_s-1}$. Denote by $t_1,\ldots,t_{n-n_s}$ the homogeneous coordinates of $\mathbb{P}^{n-n_s-1}$ and let $Y_s<X_s\times\mathbb{P}^{n-n_s-1}=E$ be the subvariety defined by $t_1,\ldots,t_{n-d_s-1}$. For $i=1,\ldots,s-1$ let Y_i be the subvariety of $X_i\times\mathbb{P}^{n-n_s-1}$ defined by $t_1,\ldots,t_{n-n_s-1}$. Then $Y_i\cong X_i$ for $i\neq s$ and we have a chain of smooth T-varieties $Y=Y_{s+1}>Y_s>\cdots>Y_1$. Let y_i be the generic point of Y_i. Then $y_{s+1},\ldots,y_s$ is a T-central chain, $f(y_i)=x_i$, $\dim y_i=\dim x_i$ for $i\neq s$ and $\dim y_s=d_s$.

§ 3. Valuations and algebraic limits

Once again, we assume that $\mathcal{T}$, T and k are given as in § 1.

(3. 1) Definition. Let F/k be a function field, $F\in T$. The intersection of all T-valuation rings B of F/k is called the ring of the holomorphic functions (with respect to T) and denoted by $H=H(F/k, T)$.

If T is the class of all fields of char 0, then $H=$ algebraic closure of k in F. The ring H was studied in detail in the cases where T is the class of all formally p-adic fields (of fixed finite p-rank d) [J-R] or the class of all formally real fields [S 3]. Many facts from these special cases carry over to our general situation.

(3. 2) Definition. Let V be a k-variety. The ring $R(V):=\bigcap_{x\in V(k)} \mathcal{O}_{(x,V)}$ is called the ring of the regular functions on V.

(3. 3) Proposition. *If V is a quasiprojective k-variety and k not algebraically closed, then there exists an open affine subvariety $U \subset V$ with $V(k) \subset U$. Therefore $R(V) = S^{-1}k[U]$ where*

$$S = \{s \in k[U] \mid s(u) \neq 0 \quad \text{for all} \quad u \in U(k)\}.$$

Moreover the prime ideals of $R(V)$ correspond to the points $x \in U$ with $\bar{x} \cap U(k) \neq \emptyset$.

Proof. We have $V \subset \mathbb{P}_k^n$ for some $n \in \mathbb{N}$. We may find a form $c \in k[X_1, \ldots, X_n]$ which has no nontrivial zero in k^n, since k is not algebraically closed. Thus $V(k)$ lies in the complement U of the zero-set of c which is affine. For $h \in R(V)$ we have $h = p/q$, $p, q \in k[U]$ and the ideal $A = \{q \in k[U] \mid qh \in k[U]\}$ has no zeros in $U(k)$. Suppose $A = (q_1, \ldots, q_m)$ and $d \in k[X_1, \ldots, X_m]$ is again a form without zeros in k^m. Then $d(q_1, \ldots, q_m) \in A \cap S$.

Finally assume that $\bar{x} \cap U(k) \neq \emptyset$. Then clearly $S \cap x = \emptyset$. On the other hand, if $\bar{x} \cap U(k) = \emptyset$ and x as an ideal of $k[U]$ is generated by say $g_1, \ldots, g_m$ then $d(g_1, \ldots, g_m) \in S \cap x$.

(3. 4) Remark. Suppose that k is existentially closed in T. Let V be a k-variety with $F = k(V)$. Then $R(V) = \bigcap_{x \in V(T)} \mathcal{O}_{(x,V)}$. If, moreover, V is complete, then $R(V) \subset H(F/k, T)$.

Now let X be a fixed projective k-variety. We consider the class $B(X)$ of all pairs (Y, f), Y a projective k-variety and $f\colon Y \to X$ a birational morphism. $B(X)$ is a directed system with respect to the factorization:

$(Y_1, f_1) \geqq (Y, f) \Leftrightarrow$ there exists a birational morphism $h\colon Y_1 \to Y$ with $f_1 = f \circ h$.

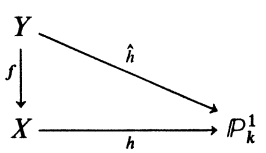
We set $F = k(X)$ which will be identified with $k(Y)$ by f^* for all $(Y, f) \in B(X)$.

This yields an inductive system of imbeddings $f^{*-1}\colon R(Y) \to F$.

(3. 5) Proposition. *Let k be existentially closed in T. Then*

$$H = H(F/k, T) = \varinjlim R(Y).$$

Proof. We saw already in (3. 4) that "$\supset$" holds. Now suppose that $h \in H$. We consider h as a rational map $X \to \mathbb{A}_k^1 \to \mathbb{P}_k^1$. By Hironakas theorem on the elimination of points of indeterminacy ([H], Main Theorem II) there exists $(Y, f) \in B(X)$, Y smooth, and a morphism $\hat{h}\colon Y \to \mathbb{P}_k^1$ such that the diagram



commutes. We have $Y(k) = Y\{k\}$ by (2. 1). Then for all $y \in Y(k)$ we get $h = p/q$, $p, q \in \mathcal{O}_{(y,Y)}$ not both in $\mathcal{M}_{(y,Y)}$ and, since y is the centre of a T-place, $q \notin \mathcal{M}_{(y,Y)}$ which yields the assertion.

(3. 6) Corollary. *If k is existentially but not algebraically closed, then $H = H(F/k, T)$ is a Prüfer domain with field of fractions F.*

Proof. For a maximal ideal m of H we show that H_m is a valuation ring of F. So for $h \in F = k(X)$ we choose (Y, f) and $\hat{h}$ as in the preceding proof. By (3. 3) and (3. 4) there exists an open affine subvariety U of Y with $Y(k) \subset U$ and $k[U] \subset R(Y) \subset H$. Let y be a maximal ideal of $k[U]$ with $y \supset m \cap k[U]$. We have $h = p/q$, $p, q \in k[U]$, not both in y. Hence $h \in \mathcal{O}_{(y,Y)}$ or $h^{-1} \in \mathcal{O}_{(y,Y)}$ but $\mathcal{O}_{(y,Y)} \subset H_m$. q.e.d.

For a Prüfer domain H the class group $\mathrm{Cl}(H)$ is defined to be the group $I(H)$ of all invertible (= finitely generated) fractional ideals of H modulo the subgroup of the fractional principal ideals. If $\mathrm{Cl}(H)$ is trivial, H is called a Bezout ring.

The following is a consequence of a result of Roquette ([R], Theorem 1).

(3. 7) Proposition. *Let k be existentially closed in T but not algebraically closed and let F/k be a function field.*

Then $\mathrm{Cl}(H)$ has finite exponent for $H = H(F/k, T)$. Moreover, if there exist irreducible polynomials $f, g \in k[X]$ with mutually prime degrees, then H is a Bezout-Ring.

In order to approximate $I(H)$ and $\mathrm{Cl}(H)$ we assume that our fixed projective k-variety X is smooth. Furthermore we consider the cofinal subclass $B'(X)$ consisting of all pairs $(Y, f) \in B(X)$ where Y is smooth ([H], Main Theorem I). As usual a prime divisor of Y will be a closed subvariety of Y of codimension 1. There are two possibilities to define T-divisors.

1. The prime divisor P is called a weak T-prime divisor, if P contains a T-point.

2. The prime divisor P is called a strong T-prime divisor, if P is a T-variety.

The weak T-prime divisors generate a subgroup $\mathrm{Div}_T(Y)$ in the group $\mathrm{Div}(Y)$ of all Weil-divisors of Y and we have a restriction map

$$\mathrm{res}_T \colon \mathrm{Div}(Y) \longrightarrow \mathrm{Div}_T(Y).$$

By $\mathrm{Cl}_T(Y)$ we denote the weak T-classgroup of Y, that is the group $\mathrm{Div}_T(Y)/\mathrm{res}_T P(Y)$ where $P(Y)$ is the group of the principal divisors of Y.

Denoting again by $I(R(Y))$ the group of the invertible ideals of $R(Y)$ we get from (3. 3) and the fact, that $R(Y)$ is locally factorial, that

$$\mathrm{Div}_T(Y) \cong I(R(Y)) \quad \text{and} \quad \mathrm{Cl}_T(Y) \cong \mathrm{Cl}(R(Y)),$$

if k is existentially closed.

The strong T-prime divisors generate a subgroup $\mathrm{Div}^T(Y)$ in $\mathrm{Div}(Y)$. We define the strong T-class group $\mathrm{Cl}^T(Y)$ correspondingly.

Then $\mathrm{Div}^T(Y)$ is no longer isomorphic to $I(R(Y))$ in a canonical way, but we still have

(3. 8) Proposition. *If k is not algebraically closed but existentially closed in T there are canonical isomorphisms*

$$\varinjlim \mathrm{Div}_T(Y) \cong I(H(F/k, T)) \cong \varinjlim \mathrm{Div}^T(Y),$$
$$\varinjlim \mathrm{Cl}_T(Y) \cong \mathrm{Cl}(H(F/k, T)) \cong \varinjlim \mathrm{Cl}^T(Y).$$

Here for a birational morphism $g: Y' \to Y$ of smooth projective k varieties $g^{*T}: \mathrm{Div}^T(Y) \to \mathrm{Div}^T(Y')$ is defined by

$$g^{*T} := \mathrm{res} \mid \mathrm{Div}^T(Y') \circ g^* \quad \text{where} \quad g^*: \mathrm{Div}(Y) \to \mathrm{Div}(Y')$$

is the natural pull back.

Proof. We get the isomorphism at the left hand side by (3. 5) and commutation with the formation of the direct limit. Concerning the right hand isomorphisms we have for each $(Y, f) \in B'(X)$ the map

$$h_Y: \mathrm{Div}^T(Y) \to I(H(F/k, T))$$

since $\mathrm{Div}^T(Y) \subset \mathrm{Div}_T(Y)$. Now consider $(Y', f') \in B'(X)$ with a birational morphism $g: Y' \to Y$ such that $f' = f \circ g$. We claim that $h_Y = h_{Y'} \circ g^{*T}$. It is clear that $h_Y(D) \subset h_{Y'} \circ g^{*T}(D)$ for all $D \in \mathrm{Div}^T(Y)$. To prove the other implication let $p: Z \to Y$ be a birational morphism which is given by a finite sequence of monoidal transformations at smooth centres such that there exists a birational morphism $q: Z \to Y'$ with $p = g \circ q$ ([H], p. 144). Then for all $D \in \mathrm{Div}^T(Y)$ we have $h_Y(D) = h_Z \circ p^{*T}(D)$ which proves the claim. This yields a homomorphism

$$h: \varinjlim \mathrm{Div}^T(Y) \to I(H(F/k, T)).$$

On the other hand for $J \in I(H(F/k, T))$, say $J = (a_1, \ldots, a_m)$, consider $(Y, f) \in B'(X)$ such that the a_i admit no points of indeterminacy and the supports of the divisors of the a_i have smooth components ([H], Main Theorem II, Corollary 1). Then $J = h_Y(D)$ for a $D \in \mathrm{Div}^T(Y)$. Now the map $d: J \mapsto$ image of D in $\varinjlim \mathrm{Div}^T(Y)$ is well defined and inverse to h.

From this one easily gets also the isomorphism

$$\mathrm{Cl}(H(F/k, T)) \cong \varinjlim \mathrm{Cl}^T(Y).$$

Finally we describe the set of all T-valuations of F by an algebraic limit.[5]) Assume, that the projective model X of F and the inverse system $B(X)$ is given as before. Let $B(X, T)$ be the inverse system which we get by restriction of the elements (Y, f) of $B(X)$ to $(Y(T), f)$. The elements of $\varprojlim Y(T)$ will be represented by families $\{y_\lambda\}$, $y_\lambda \in (Y_\lambda(T), f_\lambda)$ with $g_{\lambda\mu}(y_\lambda) = y_\mu$ for $g_{\lambda\mu}: Y_\lambda \to Y_\mu$ with $f_\lambda = f_\mu \circ g_{\lambda\mu}$.

Now let $\mathrm{val}(F/k, T)$ be the Zariski space of all T-valuations of F/k and for $v \in \mathrm{val}(F/k, T)$ let B_v be the corresponding valuation ring of F.

---

[5]) This idea already appears in Zariski's work, see [Z], § 9.

(3. 9) Proposition. *Suppose that the class T is closed under direct limits. Then the maps*

$$\mathrm{val}(F/k, T) \underset{b}{\overset{z}{\rightleftarrows}} \varprojlim Y(T);$$

$$v \xrightarrow{z} \textit{system of centres of } v,$$

$$\{y_\lambda\} \xrightarrow[b]{} \textit{valuation } v \textit{ with } B_v = \varinjlim \mathcal{O}_{(y_\lambda, Y_\lambda)}$$

are mutually inverse homeomorphisms. Moreover, for the residue field of $v = b\{y_\lambda\}$ *one has* $F_v \cong \varinjlim k(y_\lambda)$.

Proof. The crucial point is to show that $\varinjlim \mathcal{O}_{(y_\lambda, Y_\lambda)}$ is a valuation ring in F. In fact again this follows directly from Hironakas theorem on the elimination of points of indeterminacy.

The valuegroup $|F|_v$ for $v = b\{y_\lambda\}$ can also be described. Denoting again for a subring $A \subset F$ by $I(A)$ the group of the invertible fractional ideals of A we have $|F|_v = I(B_v)$. Commuting the formation of $I(\mathcal{O}_{(y_\lambda, Y_\lambda)})$ with the direct limit we get

(3. 10) Proposition. *Under the above notations let* $v = b\{y_\lambda\}$. *There is a canonical isomorphism* $|F|_v \cong \varinjlim I(\mathcal{O}_{(y_\lambda, Y_\lambda)})$.

Recall that $I(\mathcal{O}_{(y_\lambda, Y_\lambda)}) \cong \mathrm{Div}(\mathcal{O}_{(y_\lambda, Y_\lambda)})$ if y_λ is simple.

Let Y be smooth. With the notation of (3. 10) the group $I(\mathcal{O}_{(y, Y)})$ may be considered as the local divisor group of Y in y, generated by the minimal prime ideals of $\mathcal{O}_{(y, Y)}$ which are T-prime divisors in the weak sense. Obviously, the canonical map $I(\mathcal{O}_{(y, Y)}) \to |F|_v$ is surjective. We show that this holds true even for the strong local T-divisor group $\mathrm{Div}^T(\mathcal{O}_{(y, Y)})$, which is generated by the strong T-prime divisors.

(3. 11) Proposition. *Let Y be a smooth variety with function field F and let v be a T-valuation with centre y on Y. Then the natural map*

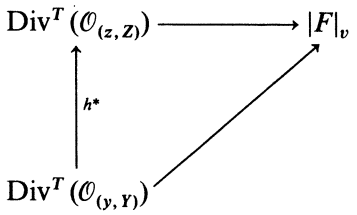
$$f_v \colon \mathrm{Div}^T(\mathcal{O}_{(y, Y)}) \to |F|_v$$

is surjective.

Proof. For a prime divisor $P \in \mathrm{Div}^T(\mathcal{O}_{(y, Y)})$ and a local parameter u of P in y we have $f_v(P) = v(u)$.

Given $\gamma \in |F|_v$ choose a function $g \in F^*$ with $v(g) = \gamma$. There is a composition of monoidal transformations with smooth irreducible centers $h \colon Z \to Y$, Z smooth, such that the divisor $\mathrm{div}(g)$ of g on Z consists of smooth prime divisors ([H], Main Theorem II, Corollary 1). Let z be the centre of v on Z and $\mathrm{div}_z(g) = \prod_{i=1}^{k} C_i^{r_i}$ be the

restriction of $\operatorname{div}(g)$ to $\operatorname{Div}(\mathcal{O}_{(z,Z)})$. Since the prime divisors C_i contain a simple T-point, they are strong T-prime divisors by (2. 3). Therefore, $\operatorname{div}_z(g) \in \operatorname{Div}^T(\mathcal{O}_{(z,Z)})$ and $f_v(\operatorname{div}_z(g)) = v(g)$. Now assume that $h: Z \to Y$ is a monoidal transformation with the smooth irreducible centre W. Then one obtains a commutative diagram



It remains to show that h^* is surjective: Let $E = h^{-1}(W)$ be the exceptional divisor. If $y \notin W$ or equivalently $z \notin E$ there is nothing to prove. So assume that $y \in W$. For a prime divisor $P \in \operatorname{Div}^T(\mathcal{O}_{(y,Y)})$ we have $h^*(P) = \tilde{P}^t E^l$ where $\tilde{P}$ is the strict transform of P, $t = 0$ if $z \notin \tilde{P}$, $t = 1$ if $z \in \tilde{P}$ and l is the order of P at the generic point w of W. Since W is smooth and $y \in W$ is a T-point we get by (2. 3) that w is a T-point. Choose $u_1, \ldots, u_s \in \mathcal{M}_{(y,Y)} \setminus \mathcal{M}^2_{(y,Y)}$ such that they define a system of local parameters in $\mathcal{O}_{(w,Y)}$. We may assume, that $s \geqq 2$. Then u_i defines a T-prime divisor $P_i \in \operatorname{Div}^T(\mathcal{O}_{(y,Y)})$ for $i = 1, \ldots, s$. Now $z \notin \tilde{P}_i$, $\tilde{P}_i$ = strict transform of P_i, for at least one $i \in 1, \ldots, s$. In fact locally Z can be considered as a subvariety of $Y \times \mathbb{P}^{s-1}$, namely

$$Z = \{(x, [t_1, \ldots, t_s]) \mid u_i(x) t_j = u_j(x) t_i \quad \text{for} \quad 1 \leqq i, j \leqq s\}.$$

Assume $t_i(z) \neq 0$. Then $t_i = 0$ on $\tilde{P}_i \setminus E$, hence $t_i = 0$ on $\tilde{P}_i$ and $z \notin \tilde{P}_i$. Therefore $h^*(P_i) = E$ for at least one P_i which proves the surjectivity.

(3. 12) Examples. Let T be the class of the formally real fields. Consider $X = \mathbb{P}^2_R$. We have $\mathbb{A}^2_R \subset \mathbb{P}^2_R$. Set $x = (0, 0) \in \mathbb{A}^2_R$. We are interested in real valuations v of $F = R(X_1, X_2)$ with centre x.

1) v is discrete of rank 1 and $\dim v = 1$: consider the chain $\{Y_n, f_n\} \subset B(X)$ with $Y_0 = X$ and Y_{n+1} = blowing up of Y_n at the centre of v. Then at a number $m \in \mathbb{N}$ it happens, that the centre of v in Y_m is a closed point but the centre of v in Y_{m+1} is the exceptional divisor E. Here the chain $\{Y_n, f_n\}$ becomes stationary and for $e = x(E)$ we get $B_\varphi = \mathcal{O}_{(e, Y_{m+1})}$. In particular $F_v \cong R(t)$ is purely transcendental.

2) v is discrete of rank 2 hence $F_v \cong R$. Then there exists a real valuation w, coarser than v, which is discrete of rank one, hence $\dim w = 1$. By 1) the centre of v is a closed real point on the centre of w, which is a real divisor for all (Y_λ, f_λ) above a suitable $(Y_{\lambda_0}, f_{\lambda_0})$.

3) v is discrete of rank 1 and $\dim v = 0$, that is $F_v = R$. As a special example, for $p(X_1, X_2) \in R(\mathbb{A}^2)$ set $|p|_v :=$ subdegree of $p(X_1, \exp(X_1) - 1)$. Let us form the chain $\{Y_n, f_n\}$ as in 1). Then the lifting of the curve $\{(t, e^t - 1) \mid t \in R\}$ in Y_n intersects the exceptional divisor E_n of Y_n at the centre y_n of v. Consider the blowing up $g_{n+1}: Y_{n+1} \to Y_n$. For a prime-divisor $P \in \operatorname{Div}(Y_n)$ let $\tilde{P} < Y_{n+1}$ be its strict transform.

Then

$$g_{n+1}^*: \mathrm{Div}(\mathcal{O}_{(y_n, Y_n)}) \to \mathrm{Div}(\mathcal{O}_{(y_{n+1}, Y_{n+1})})$$

maps P to $\tilde{P}^t \cdot E_{n+1}^l$ where $l =$ order of P at y_n, $t=0$ if $y_{n+1} \notin \tilde{P}$ and $t=1$ if $y_{n+1} \in \tilde{P}$. Hence for a prime-divisor P of X and sufficiently large $n \in \mathbb{N}$ one has

$$g_n^* \circ g_{n-1}^* \circ \cdots \circ g_1^*(P) = E_n^k$$

where k is the order of which P "touches" the curve $\{(t, e^t - 1) \mid t \in \mathbb{R}\}$ in $x = y_0$.

This shows directly: $\varinjlim \mathrm{Div}(\mathcal{O}_{(y_\lambda, Y_\lambda)}) \cong \mathbb{Z} \cong |F|_v$.

4) rank $v=1$, $\dim v=0$ and $\dim \mathbb{Q} \otimes |F|_v = 2$. As a special example for this take $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ say $\alpha > 1$. The substitution $X_1 \to t$, $X_2 \to t^\alpha$ imbeds $F = \mathbb{R}(X_1, X_2)$ into the field $\mathbb{R}((G))$ of Hahn power series with exponents in $G = \mathbb{Z} + \alpha\mathbb{Z} \subset (\mathbb{R}, +)$. Now let v be the pull back of the natural valuation of $\mathbb{R}((G))$. Again let us consider the chain (Y_n, f_n) as above where $y_n =$ centre of v in Y_n and $E_n =$ exceptional divisor of $g_n: Y_n \to Y_{n-1}$. For convenience we denote a prime-divisor in Y_{n-1} and its strict transform in Y_n by the same letter.

Now let $\{a_0, a_1, a_2, \ldots\}$, $a_i \in \mathbb{N}$, be the sequence of continued fractions of α and set $b_n := \sum_{i=0}^{n} a_i$. Finally set P_1, P_2 for the varieties of X_1 and X_2 in $X = \mathbb{P}_{\mathbb{R}}^2$. Then the sequence $\{y_n\}$ of centres is represented by transversal intersections of prime-divisors as follows:

$$\underbrace{P_2 \cap P_1,\ P_2 \cap E_1, \ldots, P_2 \cap E_{b_0}}_{a_0},\ \underbrace{E_{b_0} \cap E_{b_0+1}, \ldots, E_{b_0} \cap E_{b_1}}_{a_1},$$

$$\underbrace{E_{b_1} \cap E_{b_1+1}, \ldots, E_{b_1} \cap E_{b_2}, \ldots}_{a_2}.$$

At the point y_j for $b_n < j < b_{n+1}$ one has $g_{j+1}^* E_j = E_{j+1}$. Hence the elements of $\varinjlim \mathrm{Div}\, \mathcal{O}_{(y_n, Y_n)}$ correspond to sequences $\mathbb{N} \to \mathbb{Z}^2$, $j \to (c_j, d_j)$ such that

$$(c_{j+1}, d_{j+1}) = \left.\begin{array}{ll} (c_j, d_j + c_j) & \text{for } j \neq b_n \\ (d_j, c_j + d_j) & \text{for } j = b_n \end{array}\right\} \text{ for sufficiently large } j.$$

In order to get back this group, which is the value group of v, as a subgroup of $\mathbb{R}$ let the sequence $\{(r_j, s_j)\}$ correspond to the element, which is represented by P_1 in $\varinjlim \mathrm{Div}\, \mathcal{O}_{(y_n, Y_n)}$. Then one takes the map

$$\varphi: \{(c_j, d_j)\} \to \lim_{n \to \infty} d_{b_n}/s_{b_n} \in \mathbb{R}.$$

Of course, $\varphi(X_1) = 1$ and by elementary formulas for continued fractions $\varphi(X_2) = \alpha$. On the other hand each element of $|F|_v$ is of the form $v(X_1^m X_2^n)$ for $m, n \in \mathbb{Z}$. Therefore by (3. 10) and (3. 11) each sequence $\{(c_j, d_j)\}$ of the above form comes from some $X_1^m X_2^n$. It follows, that $\lim_{n \to \infty} d_{b_n}/s_{b_n}$ always exists and φ defines an isomorphism:

$$\varinjlim \mathrm{Div}^T_{(y_n, Y_n)} \to |F|_v \subset \mathbb{R}.$$

Finally we deal with the question, for which subclasses D of T-valuations of F/k one has $H(F/k, T) = \bigcap_{v \in D} B_v$. This was done in [S2] in the case where $k = \mathbb{R}$ and T is the class of the formally real fields. The problem was also considered in [K-P] where nice classes D are presented. In the following theorem we classify all possibilities for D. The motivation for this question and related ones will be given in § 4.

(3. 13) Theorem. *Let D be a set of T-valuations of F/k, where k is existentially but not algebraically closed. The following conditions are equivalent:*

a) $H(F/k, T) = \bigcap_{v \in D} B_v$.

b) *For every quasi-projective variety X over k with $k(X) = F$ each rational simple point $x \in X(k)$ is centre of a valuation $v \in D$.*

Here the implication b) $\rightarrow$ a) *holds also true if k is algebraically closed.*

Proof. b) $\rightarrow$ a). Given $f \in \bigcap_{v \in D} B_v$ choose a smooth projective k-variety X with $F = k(X)$ such that f defines a morphism: $X \rightarrow \mathbb{P}^1_k$ (again we use Hironakas result on the elimination of points of indeterminacy). Let $x \in X$ be a rational point and $v \in D$ with centre x. Since $f \in B_v$ we have $f(x) \in \mathbb{A}^1_k$, hence $f \in \mathcal{O}_{(x,X)}$. This shows that $f \in R(X)$ but $R(X) \subset H(F/k, T)$ by (3. 4).

a) $\rightarrow$ b). Assume, there is a quasi-projective k-variety X with $k(X) = F$ and a simple rational point $x \in X$ which is not centre of a valuation $v \in D$. Since k is not algebraically closed, we find an imbedding $X \hookrightarrow \mathbb{P}^m_k$ such that $X(k) \subset \mathbb{A}^m_k \subset \mathbb{P}^m_k$ (compare (3. 3)). Represent x by coordinates $(a_1, \ldots, a_m)$ in $\mathbb{A}^m_k$. As earlier we choose a form $\varphi \in k[X_1, \ldots, X_m]$ without nontrivial zeros in k^m. Then the function

$$h := \varphi(X_1 - x_1, \ldots, X_m - x_m)$$

satisfies $h \in \mathcal{M}_{(x,X)}$ and $h \in \mathcal{O}^*_{(y,X)}$ for all $y \in X(k)$, $y \neq x$. Let z be the centre of an arbitrary valuation $v \in D$. Then $\bar{z} \cap X(k) \neq \{x\}$, since z is a T-point and k is existentially closed. Therefore $h \in \mathcal{O}^*_{(z,X)}$. Now $h^{-1} \in \bigcap_{v \in D} B_v$ but $h^{-1} \notin H(F/k, T)$.

§ 4. Application

In this section we shall apply our results to Becker's study of n-th powers. In ([B], 1. 9) he proves

Theorem. *Let K be a formally real field, $a \in K$. Then a is a sum of $2m$-th powers if and only if a is a sum of squares and $v(a) \in 2m \cdot |K|_v$ for all valuations v of K with a formally real residue field.*

If K is a function field over $\mathbb{R}$, then every valuation of K with formally real residue field is trivial on $\mathbb{R}$. The theorem now leads to the following definition.

(4. 1) Definition. Let K/k be a function field. A function $f \in K$ is called n-divisible (with respect to T) if and only if $v(f) \in n|K|_v$ for every T-valuation v of K over k.

In case $\mathcal{T}$ is the theory of formally real fields and $k = \mathbb{R}$ the following is proved in [S2].

(*) *$f \in K$ is n-divisible if and only if $n|v(f)$ for every valuation v of K over k having a formally real residue field K_v with $d(K_v/k) = d(K/k) - 1$.*

In [P] Prestel proves that in case k is an archimedean ordered field and F is the field of rational functions $F = k(X_1, \ldots, X_m)$ the valuations induced by power series are sufficient. In this chapter we shall generalize (*) and deduce a characterization of n-divisible functions by means of rational valuations with value group $\mathbb{Z}$. We assume the axioms A0, A1, A2.

(4. 2) Theorem. *Let k be existentially closed and K/k be a function field, $K \in T$ and f a fixed element $\in K$. There are finitely many discrete rank one T-valuations $v_1, \ldots, v_r$ with $d(K_{v_i}/k) = d(K/k) - 1$ (and value group $|K|_{v_i} = \mathbb{Z}$) such that for every T-valuation v over k and every $n \in \mathbb{N}$: $v(f) \notin nK_v$ implies $v_i(f) \notin n|K|_{v_i}$ for at least one $i \in \{1, \ldots, r\}$. More precisely, $v(f) = \sum_{i=1}^{r} v_i(f) r_i$, $r_i \in |K|_v$. The valuations v_i are induced by 1-codimensional smooth T-subvarieties K_i of a smooth variety Y over k with function field K.*

Proof. Choose a projective smooth variety X over k with function field K. Let $\operatorname{div}_X(f) = C_1^{r_1} \cdots C_k^{r_k}$ be the divisor of f on X, $C_1, \ldots, C_r$ are subvarieties of codimension 1 of X. Hironaka's main theorem II [H] implies that we can find a proper morphism $h: Y \longrightarrow X$ which is a composition of monoidal transformations in smooth centres, such that the preimage $h^{-1}(C_1 \cup \cdots \cup C_k)$ is the union of 1-codimensional smooth subvarieties $K_1, \ldots, K_s$ of Y. In particular Y is projective and smooth and h induces an isomorphism between the function fields $k(X)$ and $k(Y)$. The divisor of f on Y is $\operatorname{div}_Y(f) = K_1^{t_1} \cdots K_s^{t_s}$, $t_i \in \mathbb{Z}$. For let E be a 1-codimensional subvariety of Y with $\operatorname{ord}_E(f) \neq 0$, where ord_E denotes the discrete valuation induced by E. Let x be a closed point of E. Then $f \notin \mathcal{O}^*_{(x,Y)}$ hence $f \notin \mathcal{O}^*_{(h(x),X)}$. This is only possible if $h(x) \in C_1 \cup \cdots \cup C_k$ ([M], Theorem 38).

We may assume that $K_1, \ldots, K_r$, $r \leqq s$, are T-varieties and K_i is not a T-variety for $i > r$. Then $K_1, \ldots, K_r$ induce T-valuations v_i of codimension $= 1$. Now let v be a T-valuation over k and y its center on Y. Then $y \notin K_i$ for $i > r$ by (2. 3). In $\mathcal{O}_{(y,Y)}$ let K_i be defined by u_i. So $u_i \in \mathcal{O}^*_{(y,Y)}$ for $y \notin K_i$, and u_i is a prime element otherwise. Now

$$f = \prod_{i=1}^{r} u_i^{t_i} u_0, \quad t_i \in \mathbb{N}, \quad u_0 \in \mathcal{O}^*_{(y,Y)},$$

thus

$$v(f) = \sum_{i=1}^{r} t_i v_i(u_i) = \sum_{i=1}^{r} w_i(f)\, v_i(u_i). \qquad \text{q.e.d.}$$

From this theorem one easily derives a characterization of the n-divisible functions by means of valuations with value group $\mathbb{Z}$ and residue field k.

(4. 3) Corollary. *The assumptions being the same as in the theorem, there exist finitely many valuations $w_1, \ldots, w_r$, all rational over k with value group $\mathbb{Z}$ such that for every T-valuation v of K/k one has $v(f) = \sum_{i=1}^{r} w_i(f) r_i$ with $r_i \in |K|_v$.*

Proof. We use the same notations as in the preceding proof. Fix $i \in \{1, \ldots, r\}$ and choose $x \in K_i(k)$ with $x \notin K_j$, $j \neq i$. Let U be an affine open neighbourhood of x, say $U \subset \mathcal{A}_k^m$ such that K_i is defined by the prime element $u_i = u$ of $k[U]$. We now construct a formal curve $\gamma(t)$ on U with centre x which meets K_i transversally in x, i.e. $\operatorname{ord}(u(\gamma(t))) = 1$. We may assume that U can be described by $m-n$ equations $f_{n+1} = \cdots = f_m = 0$, $f_j \in k[X_1, \ldots, X_m]$, where $n = \dim Y = d(K/k)$. Furthermore, x can be assumed to be the origin $x = 0$. As in the proof of (2. 12) we may suppose

$$\det\left(\frac{\partial(f_{n+1}, \ldots, f_m)}{\partial(X_{n+1}, \ldots, X_m)}(x)\right) \neq 0$$

and the implicit function theorem provides formal power series $\gamma_{n+1}, \ldots, \gamma_m$ in $\bar{t} := (t_1, \ldots, t_n)$ with $\gamma_j(0, \ldots, 0) = 0$ and $f_j(\bar{t}, \gamma_{n+1}(\bar{t}), \ldots, \gamma_m(\bar{t})) = 0$ in $k[[t_1, \ldots, t_n]]$, $j = n+1, \ldots, m$. Choose n power series $\gamma_1, \ldots, \gamma_n \in k[[t]]$ algebraic independent over $k(t)$ and let $\alpha_l := \left(\frac{\partial}{\partial t}\gamma_l\right)(0)$ for $1 \leqq l \leqq n$. We shall manipulate the α_l so that the curve

$$\gamma := (\gamma_1(t), \ldots, \gamma_n(t), \tilde{\gamma}_{n+1}(t), \ldots, \tilde{\gamma}_m(t))$$

with $\tilde{\gamma}_j(t) = \gamma_j(\gamma_1(t), \ldots, \gamma_n(t))$ meets K_i transversally. We have $\operatorname{ord}(u(\gamma)) = 1$ if and only if $\frac{\partial(u \circ \gamma)}{\partial t}(0) \neq 0$. Now

$$\begin{aligned}\frac{\partial(u \circ \gamma)}{\partial t}(0) &= \sum_{j=1}^{n} \frac{\partial u}{\partial X_j}(0) \cdot \alpha_j + \sum_{j=n+1}^{m} \frac{\partial u}{\partial X_j}(0) \cdot \sum_{l=1}^{n} \frac{\partial \gamma_j}{\partial t_l}(0) \cdot \alpha_l \\ &= \sum_{j=1}^{n} \left(\frac{\partial u}{\partial X_j}(0) + \sum_{l=n+1}^{m} \frac{\partial u}{\partial X_l}(0) \cdot \frac{\partial \gamma_l}{\partial t_j}(0)\right) \alpha_j\end{aligned}$$

implies, that one can find a suitable tupel $(\alpha_1, \ldots, \alpha_m)$ if at least one of the terms in brackets is not zero. But this follows from the fact that the derivative du is nonzero on the tangent space of Y in x. Thus we replace the coefficients of t^1 in $\gamma_1, \ldots, \gamma_n$ by $\alpha_1, \ldots, \alpha_n$. The power series γ now induces an embedding $K \to k((t))$ (see the proof of (2. 12)) and hence yields a discrete rank one valuation w_i, rational over k, $w_i(u_i) = 1$ and $w_i(u_j) = 0$ for $i \neq j$. Therefore

$$v(f) = v\left(\prod_{i=1}^{r} u_i^{t_i}\right) = \sum_{i=1}^{r} t_i v(u_i) = \sum_{i=1}^{r} w_i(f)\, v(u_i). \qquad \text{q.e.d.}$$

References

[A] *C. Andralas*, Real places in function fields. Thesis, Albuquerque 1983.

[Ar] *E. Artin*, Über die Zerlegung definiter Funktionen in Quadrate, Abh. Math. Sem. Univ. Hamburg **5** (1927), 100—115.

[B] *E. Becker*, The real holomorphy ring and sums of $2n$-th powers, in: Géométrie Algébrique Réelle et Formes Quadratiques, Proc. Rennes 1981, Lecture Notes in Math. **959**, Berlin-Heidelberg-New York 1982, 139—181.

[B1] *E. Becker*, Valuations and real places in the theory of formally real fields, in: Géométrie Algébrique Réelle et Formes Quadratiques, Proc. Rennes 1981, Lecture Notes in Math. **959**, Berlin-Heidelberg-New York 1982, 1—40.

[B-J] *E. Becker* and *B. Jacob*, Rational points on algebraic varieties over a generalized real closed field: A modul theoretic approach, J. reine angew. Math. **357** (1985), 77—95.

[Bf] *G. Brumfiel*, Partially ordered rings and semi-algebraic geometry, Cambridge 1979.

[Br] *L. Bröckner*, Real spectra and distributions of signatures, in: Géométrie Algébrique Réelle et Formes Quadratiques, Proc. Rennes 1981, Lecture Notes in Math. **959**, Berlin-Heidelberg-New York 1982, 249—272.

[Br1] *L. Bröcker*, Zur Theorie der quadratischen Formen über formal reellen Körpern, Math. Ann. **210** (1974), 233—256.

[C-K] *C. C. Chang* and *H. J. Keisler*, Model Theory, New York 1973.

[Co-CR] *M. Coste* and *M. F. Coste Roy*, La topologie du spectre réel. Ordered fields and semialgebraic geometry, Contemporary Math. **8** (1982), 27—61.

[D-K] *H. Delfs* and *M. Knebusch*, Semialgebraic topology over a real closed field. II: basic theory of semialgebraic spaces, Math. Z. **178** (1981), 175—213.

[E] *G. Efroymson*, Local reality on algebraic varieties, J. Algebra **29** (1974), 133—142.

[G] *M. J. Greenberg*, Lectures on formes in many variables, New York 1969.

[H] *H. Hironaka*, Resolution of singularities of an algebraic variety over a field of characteristic zero, Ann. Math. **79** (1964), 109—326.

[Ha] *R. Hartshorne*, Algebraic Geometry, Berlin-Heidelberg-New York 1977.

[J-R] *M. Jarden* and *P. Roquette*, The Nullstellensatz over p-adically closed fields, J. Math. Soc. Japan **32** (1980), 425—460.

[K-P] *F. V. Kuhlmann* and *A. Prestel*, On places of algebraic function fields, J. reine angew. Math. **353** (1984), 181—195.

[L] *T. Y. Lam*, Orderings, valuations and quadratic forms, Regional conference series in mathematics 52, Providence, Rhode Island 1983.

[M] *H. Matsumura*, Commutative Algebra, New York 1970.

[M-S] *S. Maclane* and *O. F. G. Schilling*, Zero-dimensional branches of rank one on algebraic varieties, Ann. Math. **40** (1936), 507—520.

[P] *A. Prestel*, Model theory of fields: an application to positive semidefinite polynomials, preprint.

[P-R] *A. Prestel* and *P. Roquette*, Formally *p*-adic fields, Lecture Notes in Math. **1050**, Berlin-Heidelberg-New York 1984, 92—121.

[R] *P. Roquette*, Principal ideal theorem for holomorphy rings in fields, J. reine angew. Math. **262/263** (1973), 361—374.

[S1] *H. W. Schülting*, Real holomorphy rings in real algebraic geometry, in: Géométrie Algébrique Réelle et Formes Quadratiques, Proc. Rennes 1981, Lecture Notes in Math. **959**, Berlin-Heidelberg-New York 1982, 433—442.

[S2] *H. W. Schülting*, Prime divisors on real varieties and valuation theory, to appear in J. of Algebra.

[S3] *H. W. Schülting*, On real places and their holomorphy ring, Comm. Alg. **10** (1982), 1239—1284.

[Z-S] *O. Zariski* and *P. Samuel*, Commutative Algebra. II, Berlin-Heidelberg-New York 1960.

[Z] *O. Zariski*, Applicazioni geometriche della teoria delle valutazioni, Rendiconti di Matematica e delle sue applicazioni (5) **13**, Roma 1954.

Universität Münster, Mathematisches Institut, Einsteinstraße 62, D-4400 Münster

Mathematisches Institut, Universität Dortmund, Postfach 500 500, D-4600 Dortmund 50

Eingegangen 26. Oktober 1984

Liaison of monomial curves in $\mathbb{P}^3$

By *H. Bresinsky* at Orono and *C. Huneke**) at West Lafayette

1. Introduction

In this paper we are interested in solving the following problem: let C_1 and C_2 be two monomial curves in $\mathbb{P}_k^3$. When are C_1 and C_2 in the same even liaison class? We always take "curve" to mean a closed subscheme C of $\mathbb{P}_k^3$ which is generically a complete intersection, unmixed, equidimensional and one dimensional. A "monomial curve" we will take to mean a curve embedded in $\mathbb{P}_k^3$ with generic zero

$$(t_0^{n_3}, t_0^{n_3-n_1}t_1^{n_1}, t_0^{n_3-n_2}t_1^{n_2}, t_1^{n_3})$$

where $n_1 < n_2 < n_3$ are positive integers with $(n_1, n_2, n_3) = 1$. In particular any monomial curve is irreducible and reduced.

The Hartshorne-Rao module, $M(C)$, defined to be

$$\bigoplus_{n\in Z} H^1(\mathbb{P}^3, \mathcal{I}_C(n)),$$

gives a complete invariant of the liaison class of C (up to twists and k-duals) in the sense that two curves C_1 and C_2 are in the same liaison class if and only if $M(C_1) = M(C_2)$ (up to twists and k-duals) [6]. Thus our problem reduces to giving necessary and sufficient conditions for two monomial curves to have the same Hartshorne-Rao module. However, we do not solve the problem in quite this generality.

It is well known and explicitly shown in [7] that another related module may be used as an invariant, which depends upon the resolution of the homogeneous coordinate ring of C over the polynomial ring $S = k[X_0, X_1, X_2, X_3]$. Let $C \subset \mathbb{P}_k^3$ be a curve, and let $I(C)$ be the associated ideal of C, and set $R = S/I(C)$. Then R has a minimal graded resolution of the form,

$$(1.1) \qquad 0 \to \bigoplus_{j=1}^{b_3} S(-a_{3j}) \xrightarrow{\varphi_3} \bigoplus_{j=1}^{b_2} S(-a_{2j}) \xrightarrow{\varphi_2} \bigoplus_{j=1}^{b_1} S(-a_{1j}) \xrightarrow{\varphi_1} S \to R \to 0$$

where $\varphi_2 \neq 0$ and $\varphi_3 = 0$ if and only if C is arithmetically Cohen-Macaulay. Set $E(C) = \operatorname{coker} \varphi_3^*$ where by $*$ we mean $\operatorname{Hom}(\ , S)$. Then $E(C)$ is a graded S-module of

*) Partially supported by the NSF, Alfred P. Sloan Foundation.

finite dimension over k and up to twists is an invariant of the even liaison class of C (all curves linked to C in an even number of steps) while up to twists

$$E(C)^v = \mathrm{Hom}_k(E(C), k)$$

is an invariant of the odd linkage class of C. (See [7] for details.) We use $E(C)$ instead of $M(C)$ since a resolution of the form (1. 1) was given in [1] for any monomial curve. When C is a monomial curve, it happens that $b_3 = b_1 - 3$ (where b_i is the i^{th} Betti number of R as in (1. 1)). In particular since b_3 is $\mu(E(C))$ ($\mu(M) =$ least number of generators of M), b_3 and therefore $b_1 = \mu(I(C))$ are invariants of the even liaison class of C. Therefore the even liaison classes of monomial curves break up naturally into classes depending on $\mu(I(C))$. From our main Theorem (4. 1) the following hold: if $\mu(I(C)) \geqq 5$ then it is possible for there to be either infinitely or finitely many monomial curves in the even linkage class of C, and it is possible to have an arbitrarily large finite class (Example 4. 19). If $\mu(I(C)) \leqq 4$ then the even linkage class contains infinitely many monomial curves. Moreover we give an algorithm for calculating an entire even liaison class of monomial curves.

The odd liaison classes are far more difficult, as has been widely noted in general. However one can show that if $\mu(I(C)) \leqq 4$ then the even and odd liaison classes of monomial curves coincide (section 5), while if $\mu(I(C)) = 5$, then there are no monomial curves in the odd liaison class of C. In general if $\mu(I(C)) \geqq 5$ we belive this latter statement to be true.

Another related question asks which modules E can be of the form $E(C)$ for C a monomial curve.

In particular one can completely classify all such cyclic E and all E such that $(X_0, X_1, X_2, X_3)^3 E = 0$, and further give all the monomial curves C such that $E(C) = E$ for such E. A general classification seems impossible due to the arithmetic conditions which must be satisfied. For instance, any cyclic $E(C)$ with C a monomial curve must be of the form $S/(X_0^{\gamma_0}, X_1^{\gamma_1}, X_2^{\gamma_2}, X_3^{\gamma_3})$ but not all such modules can actually occur due to arithmetic conditions on the exponents γ_i.

A brief outline of the paper is as follows. We first describe the resolution in [1] for monomial curves, which is crucial for all the results in this paper. We next give a standard form for the map φ_3 in (1. 1) by choosing appropriate bases for the modules in the resolution and then show this form is uniquely determined by the module $E(C)$. (In other words there is exactly one such form in each similarity class of any matrix representation of φ_3.) We proceed to "rebuild" all possible φ_2 and φ_1 from φ_3, which gives us a parametrization of all C', monomial curves, with $E(C') = E(C)(n)$. In general terms the problem of determining the possible ways to push a map

$$0 \longrightarrow F \xrightarrow{\varphi} G$$

of free modules forward to a resolution of an ideal is an important and difficult problem. In this particular case, the resolution of monomial ideals have so much structure that we are able to completely determine the possibilities.

Finally we have an appendix which gives complete formulas for the even liaison classes of monomial curves. (See (A. 2), (A. 4), and (A. 6).) These formulas require most of this paper to derive and to prove them valid. However it is a simple process to actually apply them as we illustrate by examples. After reading the basic definitions in Section 2, the reader might skip to the appendix to obtain the complete picture of the liaison classes.

We proceed to fix our notations and definitions.

Definition 1. 2. [5] Two curves C_1 and C_2 in $\mathbb{P}_k^3$ are algebraically directly linked (which we write as $C_1 \approx C_2$) by a complete intersection X *containing* C_1, C_2 if

i) $I(C_2)/I(X) \cong \mathrm{Hom}_{\mathbb{P}^3}(\mathcal{O}_{C_1}, \mathcal{O}_X)$

ii) $I(C_1)/I(X) \cong \mathrm{Hom}_{\mathbb{P}^3}(\mathcal{O}_{C_2}, \mathcal{O}_X)$.

We observe that if C_1 is a curve contained in a complete intersection X then the subscheme C_2 defined by

$$I(C_2)/I(X) = \mathrm{Hom}_{\mathbb{P}^3}(\mathcal{O}_{C_1}, \mathcal{O}_X)$$

is directly linked to C_1. We say C and C' are linked (written $C \sim C'$) if there exist curves $C = C_0, C_1, \ldots, C_n = C'$ such that $C_i \approx C_{i+1}$ for $0 \leqq i \leqq n-1$, and say C and C' are evenly linked if n is even and oddly linked if n is odd.

Definition 1. 3. Let $n_1 < n_2 < n_3$ be three relatively prime positive integers. The monomial curve $C = C(n_1, n_2, n_3)$ is the curve whose generic zero is given by $(t_0^{n_3}, t_0^{n_3-n_1}t_1^{n_1}, t_0^{n_3-n_2}t_1^{n_2}, t_1^{n_3})$, so that $I(C) = \ker(f)$, where $f(X_0) = t_0^{n_3}$, $f(X_1) = t_0^{n_3-n_1}t_1^{n_1}$, $f(X_2) = t_0^{n_3-n_2}t_1^{n_2}$, $f(X_3) = t_1^{n_3}$. Thus, $C = \mathrm{Proj}(S/I(C))$.

2. The resolution of C and basic lemmas

In this section we briefly describe the resolution of a monomial curve $C(n_1, n_2, n_3)$. Set $I = I(C)$, and set $n = \mu(I)$. Then the resolution of

$$R = S/I (S = k[X_0, X_1, X_2, X_3])$$

always has the form [1],

$$(2.1) \qquad 0 \to \bigoplus_{j=1}^{n-3} S(-d_{3j}) \xrightarrow{\varphi_3} \bigoplus_{j=1}^{2n-4} S(-d_{2j}) \xrightarrow{\varphi_2} \bigoplus_{j=1}^{n} S(-d_{1j}) \xrightarrow{\varphi_1} S \to R \to 0.$$

By definition, $E(C) = \mathrm{coker}\, \varphi_3^*$ and we always identify $E(C)$ with any twist $E(C)(m)$ of itself.

We need to explicitly describe the maps φ_1, φ_2, φ_3 in (2. 1). We begin with a description of the map φ_1, in other words with a minimal set of generators of I. The generators may be taken to be a set of irreducible binomials. This set is obtained as follows: let $(R_{X_0})_0$ be the affine piece of C given by inverting X_0, so that $(R_{X_0})_0$ has a parametrization

$$\left(\left(\frac{t_1}{t_0}\right)^{n_1}, \left(\frac{t_1}{t_0}\right)^{n_2}, \left(\frac{t_1}{t_0}\right)^{n_3}\right).$$

Set $t=\frac{t_1}{t_0}$. All such curves are defined by at most 3 equations which may be described as follows (see [4]). Set $y_i=\frac{X_i}{X_0}$. Then the affine equations of $(R_{X_0})_0$ are generated by $\{y_i^{\alpha_i}-y_j^{\alpha_{ij}}y_k^{\alpha_{ik}}\}$ where $\{i,j,k\}=\{1,2,3\}$ and α_i is chosen to be the least positive integer such that $\alpha_i n_i \in \langle n_j, n_k\rangle$, the semigroup in $\mathbb{N}$ generated by n_j, n_k. We set

$$
\begin{aligned}
f_1 &= X_1^{\alpha_1} - X_2^{\alpha_{12}} X_3^{\alpha_{13}} X_0^{\alpha_1-\alpha_{12}-\alpha_{13}} \in I,\\
f_2 &= -X_0^{\alpha_{21}+\alpha_{23}-\alpha_2} X_2^{\alpha_2} + X_1^{\alpha_{21}} X_3^{\alpha_{23}} \in I,\\
f_3 &= -X_0^{\alpha_{31}+\alpha_{32}-\alpha_3} X_3^{\alpha_3} + X_1^{\alpha_{31}} X_2^{\alpha_{32}} \in I,
\end{aligned}
\tag{2.2}
$$

provided these polynomials are distinct. It is possible for any pair of them to be the same polynomial, but there must be at least two distinct polynomials. If there are only two such polynomials, we let f_1 be as above and define f_2 as follows:

If $f_1=f_2$ or $f_1=f_3$ we let f_2 be the other distinct polynomial of the two. If $f_2=f_3$ then $\alpha_{21}+\alpha_{23}-\alpha_2<0$ and we let

$$f_2 = X_2^{\alpha_2} - X_0^{\alpha_2-\alpha_{21}-\alpha_{23}} X_1^{\alpha_{21}} X_3^{\alpha_{23}}.$$

(In fact in this case $\alpha_{21}=0$ and furthermore $I(C)$ is a complete intersection.)

Next we define f_3 as follows:

Let $X_0^{\gamma_0}, X_1^{\gamma_1}$ be the least powers of X_0, X_1 occurring in any monomial of f_1 or f_2. If $(f_1,f_2) \nsubseteq (X_0^{\gamma_0}, X_1^{\gamma_1})$, then necessarily C itself is a complete intersection with $I(C)$ generated by f_1, f_2. This is the trivial case—all such C are arithmetically Cohen-Macaulay and by Peskine and Szpiro [5], the set of arithmetically Cohen-Macaulay curves all lie in the same liaison class. If $(f_1,f_2) \subseteq (X_0^{\gamma_0}, X_1^{\gamma_1})$, then we may write f_1 and f_2 as two of the 2×2 minors of a 2×3 matrix A_2:

$$A_2 = \begin{pmatrix} a_{11} & -a_{12} & -X_1^{\gamma_1} \\ -a_{21} & a_{22} & X_0^{\gamma_0} \end{pmatrix},$$

where $f_1 = -a_{12}X_0^{\gamma_0} + a_{22}X_1^{\gamma_1}, f_2 = -a_{11}X_0^{\gamma_0} + a_{21}X_1^{\gamma_1}$. In general, if A is an 2×3 matrix, we let $\Delta_i(A)=$ signed minor of A determined by deleting the i^{th} column of A, $1 \leqq i \leqq 3$. Then $f_1=\Delta_1(A_2)$, $f_2=\Delta_2(A_2)$. We define $f_3=\Delta_3(A_2)$. Since f_1, f_2 are binomials, it is not difficult to show as in [1] that a_{11} is a power of either X_2 or X_3 with a_{21} a power of the other X_2 or X_3, while a_{12} and a_{22} are monomials, with coefficients $+1$.

In the case that f_1, f_2, f_3 as in (2.2) are distinct, it is still possible to write $f_i=\Delta_i(A_2)$ for some matrix A_2 with monomial entries, with a_{11}, a_{21}, and a_{13} as above.

This process defines the first three generators of I (except in the case I is a complete intersection in which case two of these generators suffice).

The rest of the generators $f_4,\dots,f_n$ of I are completely determined by the following properties:

(2. 3) There are subsets

$$S_2=\{f_1,f_2,f_3\},$$
$$S_3=\{f_2,f_3,f_4\},$$
$$S_4=\{f_{\alpha(4)},f_4,f_5\},\dots,\ S_i=\{f_{\alpha(i)},f_i,f_{i+1}\}$$

of $\{f_1,\dots,f_n\}$ such that $\alpha(i)<i$, $|S_i\cap S_{i+1}|=2$, and the elements in S_i are the 2×2 minors of a 2×3 matrix A_i of the form

$$A_i=\begin{pmatrix} a_{11}^{(i)} & a_{12}^{(i)} & -a_{13}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & a_{23}^{(i)} \end{pmatrix},$$

where $a_{13}^{(i)}=X_1^{\gamma_{1i}}$, $a_{23}^{(i)}=X_0^{\gamma_{0i}}$, $a_{kj}^{(i)}$ are monomials, $f_{\alpha(i)}=\Delta_1(A_i)$, $f_i=\Delta_2(A_i)$, $f_{i+1}=\Delta_3(A_i)$, and where $X_0^{\gamma_{0i}}$ and $X_1^{\gamma_{1i}}$ are the least powers of X_0 (respectively X_1) appearing in any monomial term of $f_{\alpha(i)}$ and f_i. Moreover we require that exactly one monomial in each column of A_i have sign minus one, the other has sign plus one. Since $f_{\alpha(i)}$ and f_i have signs determined by A_{i-1}, there is a unique such assignment of signs to give $f_{\alpha(i)}=\Delta_1(A_i)$, $f_i=\Delta_2(A_i)$.

These properties determine $f_1,\dots,f_n$ completely as was shown in [1]. Furthermore this set $\{f_1,\dots,f_n\}$ has one further property we shall use:

(2. 4) The set $S_{n-1}=\{f_{\alpha(n-k)},f_{n-1},f_n\}$ is the set determined by the affine piece $(R_{X_3})_0$ in the same fashion as $\{f_1,f_2,f_3\}$ are determined from $(R_{X_0})_0$. In particular

$$\pm f_n=X_2^{\beta_2}-X_1^{\beta_{21}}X_3^{\beta_{23}}X_0^{\beta_2-\beta_{21}-\beta_{23}}.$$

The rows of each A_i determine relations on $\{f_1,\dots,f_n\}$, in fact on $\{f_{\alpha(i)},f_i,f_{i+1}\}$ since these latter elements are the signed minors of A_i. The next proposition, proved in [1], gives the map φ_2:

Proposition 2. 5. *The rows of A_i, $2\leqq i\leqq n-1$, generate all the relations on $f_1,\dots,f_n$.*

It follows that all the relations are generated by monomial relations.

Finally we describe how the map φ_3 is given.

Lemma 2. 6. *Let* $B=\begin{pmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \end{pmatrix}$ *and* $C=\begin{pmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \end{pmatrix}$ *be two 2×3 matrices with entries in S such that $\Delta_2(B)=\Delta_1(C)$ and $\Delta_3(B)=\Delta_2(C)$. Then the following sequence is a complex:*

(2. 7)
$$0\longrightarrow S\xrightarrow[(-b_4,b_1,+c_6,-c_3)]{f_3} S^4\xrightarrow[\begin{pmatrix} b_1 & b_2 & b_3 & 0\\ b_4 & b_5 & b_6 & 0\\ 0 & c_1 & c_2 & c_3\\ 0 & c_4 & c_5 & c_6\end{pmatrix}]{f_2} S^4\xrightarrow[\begin{pmatrix}\Delta_1(B)\\ \Delta_2(B)\\ \Delta_3(B)\\ \Delta_3(C)\end{pmatrix}]{f_1} S.$$

Proof. Clearly $f_2 f_1 = 0$ since the rows of a 2×3 matrix give relations on the signed minors of the matrix. We must show $f_3 f_2 = 0$. The only non-immediate calculation is to show

$$b_1 b_5 - b_2 b_4 = c_3 c_4 + c_1 c_6,$$

which holds since

$$b_1 b_5 - b_2 b_4 = \Delta_3(B) = \Delta_2(C) = -(c_1 c_6 - c_3 c_4) = c_3 c_4 - c_1 c_6,$$

and $b_1 b_6 - b_3 b_4 = -c_2 c_6 + c_3 c_5$, which holds since

$$b_1 b_6 - b_3 b_4 = -\Delta_2(B) = -\Delta_1(C) = -(c_2 c_6 - c_3 c_5) = c_3 c_5 - c_2 c_6. \qquad \square$$

It is not difficult using the exactness criterion of Buchsbaum and Eisenbud [3] to give conditions for (2.7) to be exact.

Now consider A_i and A_{i+1}. Since $|S_i \cap S_{i+1}| = 2$, and $\alpha(i) < i$, it must be true that either $f_{\alpha(i)} = f_{\alpha(i+1)}$, or $f_{\alpha(i+1)} = f_i$. In either case there is a relation on the four rows given by A_i and A_{i+1} of the form in (2.7). (Although to apply the lemma in the first case one must switch the first and last column of A_i or shift the first column of A_{i+1} to the left in the map (f_2).) The final part of the resolution is given by

Proposition 2.8 [1]. *The relations on the matrix* $[\varphi_2]$ *of Proposition* 2.5 *are generated by the relations as in* (2.7) *on the rows of* A_i *and* A_{i+1}, $1 \leqq i \leqq n-2$.

We will need one more remark concerning these resolutions.

Remark 2.9. Let $B = \begin{pmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \end{pmatrix}$, $C = \begin{pmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \end{pmatrix}$, and $D = \begin{pmatrix} d_1 & d_2 & d_3 \\ d_4 & d_5 & d_6 \end{pmatrix}$ be three 2×3 matrices with coefficients in S, and suppose $\Delta_2(B) = \Delta_1(C)$, $\Delta_3(B) = \Delta_2(C) = \Delta_1(D)$, $\Delta_3(C) = \Delta_2(D)$. Then lemma 2.6 shows the following is a complex:

$$(2.10)\quad 0 \longrightarrow S^2 \xrightarrow[\begin{pmatrix} -b_4 & b_1 & c_6 & -c_3 & 0 & 0 \\ 0 & 0 & -c_4 & c_1 & d_6 & -d_3 \end{pmatrix}]{f_3} S^6 \xrightarrow[\begin{pmatrix} B & 0 & 0 \\ & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D \end{pmatrix}]{f_2} S^5 \xrightarrow[\begin{pmatrix} \Delta_1(B) \\ \Delta_2(B) \\ \Delta_3(B) \\ \Delta_3(C) \\ \Delta_3(D) \end{pmatrix}]{f_1} S.$$

Then the 2×2 minor of f_3 given by the 3rd and 4th columns is

$$c_1 c_6 - c_3 c_4 = -\Delta_2(C) = -\Delta_3(B) = -\Delta_1(D).$$

This situation occurs in the context of the resolution of $S/I(C)$ as in Propositions 2.5 and 2.8 when there is an element in $S_i \cap S_{i+1} \cap S_{i+2}$. If so this element is necessarily either $f_{\alpha(i)}$ (when $f_{\alpha(i)} = f_{\alpha(i+1)} = f_{\alpha(i+2)}$) or f_i (when $f_i = f_{\alpha(i+1)} = f_{\alpha(i+2)}$) or f_{i+1} (when $f_{i+1} = f_{\alpha(i+2)}$). In all cases it is $f_{\alpha(i+2)} \in S_i \cap S_{i+1} \cap S_{i+2}$, and we remark that $f_{\alpha(i+2)}$ is equal to a specified 2×2 minor of the matrix $[\varphi_3]$ in the resolution of $S/I(C)$. In fact in the context of 2.10, $-\Delta_2(C) = -\Delta_3(B) = -\Delta_1(D)$ is the unique 2×2 minor of f_3 which is a binomial. Finally notice that all the entries of the last matrix $[\varphi_3]$ in the resolution of $S/I(C)$ are also entries in $[\varphi_2]$.

We will illustrate the resolutions by giving three examples. They exemplify liaison behavior which is possible for monomial curves.

Example 2. 10. Let $C=C(2, 5, 7)$. The affine piece of C with $X_0=1$ is parametrized by 2, 5, 7. The affine equations are generated by $X_3-X_1X_2$ and $X_1^5-X_2^2$, so that $f_1=X_1^5-X_2^2X_0^3$ and $f_2=X_1X_2-X_0X_3$. Write f_1 and f_2 as 2×2 minors of the matrix,

$$A_2=\begin{pmatrix} X_3 & -X_2^2X_0^2 & -X_1 \\ -X_2 & X_1^4 & X_0 \end{pmatrix},$$

and so $f_3=\Delta_3(A_2)=X_1^4X_3-X_2^3X_0^2$.

To determine f_4, we write f_2 and f_3, as 2×2 minors of

$$A_3=\begin{pmatrix} X_2^3X_0 & -X_3 & -X_1 \\ -X_1^3X_3 & X_2 & X_0 \end{pmatrix},$$

and the other 2×2 minor $\Delta_3(A_3)=-X_1^3X_3^2+X_2^4X_0$. To determine f_5 we write f_2 and $f_4\,(f_2=f_{\alpha(4)})$ as 2×2 minors of

$$A_4=\begin{pmatrix} -X_2^4 & -X_3 & -X_1 \\ X_1^2X_3^2 & X_2 & X_0 \end{pmatrix}$$

so that $f_5=-X_2^5+X_1^2X_3^3$. As f_5 is a binomial with a pure power of X_2, we obtain $I(C(2, 5, 7))=(f_1,\ldots,f_5)$, and the resolution of $I(C)$ is,

$$0 \longrightarrow S^2(-7) \xrightarrow{\varphi_3} S^6(-6) \xrightarrow{\varphi_2} \overset{S^4(-5)}{\underset{S(-2)}{\bigoplus}} \xrightarrow{\varphi_1} S$$

with

$$[\varphi_1]^t=(f_1,\ldots,f_5),$$

$$[\varphi_2]=\begin{pmatrix} X_3 & -X_2^2X_0^2 & -X_1 & 0 & 0 \\ -X_2 & X_1^4 & X_0 & 0 & 0 \\ 0 & X_2^3X_0 & -X_3 & -X_1 & 0 \\ 0 & -X_1^3X_3 & X_2 & X_0 & 0 \\ 0 & -X_2^4 & 0 & -X_3 & -X_1 \\ 0 & X_1^2X_3^2 & 0 & X_2 & X_0 \end{pmatrix},$$

and

$$[\varphi_3]=\begin{pmatrix} X_2 & X_3 & X_0 & X_1 & 0 & 0 \\ 0 & 0 & X_2 & X_3 & X_0 & X_1 \end{pmatrix}.$$

The affine piece $X_3=1$ is given by $X_0-X_1X_2$ and $X_2^5-X_1^2$ whose homogeneous equations are $X_3X_0-X_1X_2$ and $X_2^5-X_1^2X_3^3$.

The subsets S_i are,

$$S_2=\{f_1,f_2,f_3\},\quad S_3=\{f_2,f_3,f_4\},\quad S_4=\{f_2,f_4,f_5\},$$

so that $2=\alpha(4)$.

The equation $X_3X_0 - X_1X_2 \in S_2 \cap S_4$, which as it turns out (see Theorem 4. 1) is a determining factor in the liaison class of this curve. We shall show later that the set of all monomial curves linked to C is precisely the set

$$\{C(a, a+3, 2a+3) \mid (a, 3) = 1\}.$$

Example 2. 11. Let $C = C(1, 5, 8)$. The affine equations of $X_0 = 1$ are $X_1^5 - X_2$, and $X_3 - X_1^3X_2$ so that $f_1 = X_1^5 - X_2X_0^4$ and $f_2 = -X_0^3X_3 + X_1^3X_2$. We write f_1 and f_2 as 2×2 minors of the matrix

$$A_2 = \begin{pmatrix} X_3 & -X_2X_0 & -X_1^3 \\ -X_2 & X_1^2 & X_0^3 \end{pmatrix}$$

so that $f_3 = \Delta_3(A_2) = X_3X_1^2 - X_2^2X_0$. We write f_2 and f_3 as the 2×2 minors of the matrix

$$A_3 = \begin{pmatrix} X_2^2 & -X_3X_0^2 & -X_1^2 \\ -X_3 & X_1X_2 & X_0 \end{pmatrix}$$

and so $f_4 = \Delta_3(A_3) = X_1X_2^3 - X_3^2X_0^2$.

Since the least powers of X_1 and X_0 appearing in any f_i are in f_3 and f_4, it follows that $f_{\alpha(4)} = f_3$ and f_3 and f_4 are the 2×2 minors of the matrix

$$A_4 = \begin{pmatrix} X_0X_3^2 & -X_2^2 & -X_1 \\ -X_2^3 & X_3X_1 & X_0 \end{pmatrix}.$$

Then $f_5 = \Delta_3(A_4) = -X_2^5 + X_3^3X_0X_1$ has a pure power of X_2 and so $I(C) = (f_1, \ldots, f_5)$. The resolution of $I(C)$ is given by

$$0 \longrightarrow S^2(-7) \xrightarrow{\varphi_3} \begin{matrix} S^5(-6) \\ \oplus \\ S(-5) \end{matrix} \xrightarrow{\varphi_2} \begin{matrix} S^2(-5) \\ \oplus \\ S^2(-4) \\ \oplus \\ S(-3) \end{matrix} \xrightarrow{\varphi_1} S,$$

where

$$[\varphi_1] = \begin{pmatrix} X_1^5 - X_2X_0^4 \\ -X_0^3X_3 + X_1^3X_2 \\ X_3X_1^2 - X_2^2X_0 \\ -X_3^2X_0^2 + X_2^3X_1 \\ -X_2^5 + X_1X_0X_3^3 \end{pmatrix},$$

$$[\varphi_2] = \begin{pmatrix} X_3 & -X_2X_0 & -X_1^3 & 0 & 0 \\ -X_2 & X_1^2 & X_0^3 & 0 & 0 \\ 0 & X_2^2 & -X_3X_0^2 & -X_1^2 & 0 \\ 0 & -X_3 & X_1X_2 & X_0 & 0 \\ 0 & 0 & X_0X_3^2 & -X_2^2 & -X_1 \\ 0 & 0 & -X_2^3 & X_3X_1 & X_0 \end{pmatrix},$$

$$[\varphi_3] = \begin{pmatrix} X_2 & X_3 & X_0 & X_1^2 & 0 & 0 \\ 0 & 0 & X_3 & X_2^2 & X_0 & X_1 \end{pmatrix}.$$

The affine piece $X_3 = 1$ is parametrized by $3 = 8-5$, $7 = 8-1$, and 8 and has affine equations $X_2^5 - X_1 X_0$, $X_1^2 - X_2^2 X_0$, and $X_0^2 - X_2^3 X_1$ which homogenized become $-f_5 = X_2^5 - X_1 X_0 X_3^3$, $-f_4 = X_3^2 X_0^2 - X_2^3 X_1$ and $f_3 = X_3 X_1^2 - X_2^2 X_0$. In this case $S_2 = \{f_1, f_2, f_3\}$, $S_3 = \{f_2, f_3, f_4\}$, $S_4 = \{f_3, f_4, f_5\}$ and $S_2 \cap S_4 = \{f_3\}$. We shall later show that the liaison class of monomials curves containing $C(1, 5, 8)$ is just the set

$$\{C(1, 5, 8), C(3, 7, 8)\}.$$

Example 2. 12. Let $C = C(7, 18, 19)$. The affine piece $X_0 = 1$ is generated by

$$X_1^8 - X_2 X_3^2, \quad X_2^3 - X_1^5 X_3, \quad \text{and} \quad X_3^3 - X_1^3 X_2^2$$

so that $f_1 = X_1^8 - X_2 X_3^2 X_0^5$, $f_2 = -X_0^3 X_2^3 + X_1^5 X_3$, and $f_3 = -X_0^2 X_3^3 + X_1^3 X_2^2$. These binomials are the 2×2 minors of

$$A_2 = \begin{pmatrix} X_2^2 & -X_3^2 X_0^2 & -X_1^5 \\ -X_3 & X_1^3 & X_2 X_0^3 \end{pmatrix}.$$

We now write f_2 and f_3 as 2×2 minors of

$$A_3 = \begin{pmatrix} X_3^3 & -X_2^3 X_0 & -X_1^3 \\ -X_2^2 & X_1^2 X_3 & X_0^2 \end{pmatrix}$$

and so $f_4 = \Delta_3(A_3) = X_1^2 X_3^4 - X_2^5 X_0$. The least powers of X_0 and X_1 occur in f_4 so that we take $\alpha(4) = 3$ and write f_3 and f_4 as the 2×2 minors of

$$A_4 = \begin{pmatrix} X_2^5 & -X_0 X_3^3 & -X_1^2 \\ -X_3^4 & X_1 X_2^2 & X_0 \end{pmatrix}.$$

Then $f_5 = X_1 X_2^7 - X_0 X_3^7$. Now the least powers of X_1 and X_0 occur in f_5 so we take $\alpha(5) = 4$ and write f_4 and f_5 as the 2×2 minors of

$$A_5 = \begin{pmatrix} X_3^7 & -X_2^5 & -X_1 \\ -X_2^7 & X_1 X_3^4 & X_0 \end{pmatrix}$$

and this representation gives $f_6 = \Delta_3(A_5) = -X_2^{12} + X_1 X_3^{11}$. As this binomial contains a pure power of X_2 we obtain that $I(C) = (f_1, \ldots, f_6)$.

The graded resolution is given by

$$0 \longrightarrow \begin{matrix} S^2(-11) \\ \oplus \\ S(-14) \end{matrix} \xrightarrow{\varphi_3} \begin{matrix} S(-8) \\ \oplus \\ S^3(-9) \\ \oplus \\ S^2(-10) \\ \oplus \\ S^2(-13) \end{matrix} \xrightarrow{\varphi_2} \begin{matrix} S(-5) \\ \oplus \\ S^2(-6) \\ \oplus \\ S^2(-8) \\ \oplus \\ S(-12) \end{matrix} \xrightarrow{\varphi_1} S$$

where

$$[\varphi_1]=(f_1,\ldots,f_6)^t,$$

$$[\varphi_2]=\begin{pmatrix} X_2^2 & -X_3^2X_0^2 & -X_1^5 & 0 & 0 & 0\\ -X_3 & X_1^3 & X_2X_0^3 & 0 & 0 & 0\\ 0 & X_3^3 & -X_2^3X_0 & -X_1^3 & 0 & 0\\ 0 & -X_2^2 & X_1^2X_3 & X_0^2 & 0 & 0\\ 0 & 0 & X_2^5 & -X_0X_3^3 & -X_1^2 & 0\\ 0 & 0 & -X_3^4 & X_1X_2^2 & X_0 & 0\\ 0 & 0 & 0 & X_3^7 & -X_2^5 & -X_1\\ 0 & 0 & 0 & -X_2^7 & X_1X_3^4 & X_0 \end{pmatrix}$$

and

$$[\varphi_3]=\begin{pmatrix} X_3 & X_2^2 & X_0^2 & X_1^3 & 0 & 0 & 0 & 0\\ 0 & 0 & X_2^2 & X_3^3 & X_0 & X_1^2 & 0 & 0\\ 0 & 0 & 0 & 0 & X_3^4 & X_2^5 & X_0 & X_1 \end{pmatrix}.$$

We find the sets S_i are defined as $S_2=\{f_1,f_2,f_3\}$, $S_3=\{f_2,f_3,f_4\}$, $S_4=\{f_3,f_4,f_5\}$ and $S_5=\{f_4,f_5,f_6\}$. In this case $S_2\cap S_5=\emptyset$ and we will show that C is the only monomial curve in its even linkage class.

One of our main results (Theorem 4. 1) describes the possible liaison classes of monomial curves in terms of the set S_2 and S_{n-1} where $n=\mu(I(C))$. Since S_2 is the set of three generators of $I(C)$ determined from the affine piece $X_0=1$ and S_{n-1} is the set of three generators of $I(C)$ determined from the affine piece $X_3=1$, this result is very easy to apply.

3. Admissible modules

The structure of the resolution given in section two of the ideals $I(C)$ shows that the matrix $[\varphi_3]$ has a certain form which we incapsulize in a definition. This process is necessary to prove the uniqueness of this form.

Definition 3. 1. A matrix $B=(b_{ij})$ will be said to be *admissible* if

i) B is $k\times 2k+2$ for some $k\geqq 1$,

ii)

$$b_{ij}=\begin{cases} 0 & \text{for } j<2i-1,\ j>2i+2,\\ X_{l(i)}^{\gamma_{l(i)i}} & j=2i-1,\ i>1,\\ X_{m(i)}^{\gamma_{m(i)i}} & j=2i,\ i>1,\\ X_0^{\gamma_{0i}} & j=2i+1,\\ X_1^{\gamma_{1i}} & j=2i+2,\\ X_{m(2)}^{\gamma_{m(2)1}} & j=i=1,\\ X_{l(2)}^{\gamma_{l(2)1}} & i=1,\ j=2 \end{cases}$$

where $\gamma_{kj}\geqq 1$, and where $\{l(i), m(i)\}=\{2, 3\}$ for all i.

iii) $\gamma_{0i} \geqq \gamma_{0j}$, $\gamma_{1i} \geqq \gamma_{1j}$, $\gamma_{2i} \leqq \gamma_{2j}$, $\gamma_{3i} \leqq \gamma_{3j}$ whenever $i \leqq j$.

The subscript of the exponents means the following: γ_{ij} is the exponent of X_i in the j^{th} row of $[\varphi_3]$.

A visual description of B (or actually the transpose of B) is easier to understand:

$$B^t = \begin{pmatrix} X_{m(2)}^{\gamma_{m(2)1}} & 0 \ldots & 0 \\ X_{l(1)}^{\gamma_{l(2)1}} & 0 & 0 \\ X_0^{\gamma_{01}} & X_{l(2)}^{\gamma_{l(2)2}} & \vdots \\ X_1^{\gamma_{11}} & X_{m(2)}^{\gamma_{m(2)2}} & \\ 0 & X_0^{\gamma_{02}} & \\ 0 & X_1^{\gamma_{12}} & \\ \ldots & 0 \ldots & X_{l(k)}^{\gamma_{l(k)k}} \\ \ldots & \ldots & X_{m(k)}^{\gamma_{m(k)k}} \\ \ldots & \ldots & X_0^{\gamma_{0k}} \\ 0 & 0 & X_1^{\gamma_{1k}} \end{pmatrix}.$$

Finally we require one more condition:

iv) The binomial 2×2 minors of B, namely

$$Q_i = X_0^{\gamma_{0i}} X_{m(i+1)}^{\gamma_{m(i+1)i+1}} - X_{l(i+1)}^{\gamma_{l(i+1)i+1}} X_1^{\gamma_{1i}},$$

are in $I(C)$ for some monomial curve $C = C(n_1, n_2, n_3)$. We will call these minors the *distinguished* minors of B. Also to fix an ordering if $k = 1$, then B must be of the form

$$B = (X_2^{\gamma_{21}} X_3^{\gamma_{31}} X_0^{\gamma_{01}} X_1^{\gamma_{11}}).$$

We will say an S-module M is *admissible* if M has a presentation

$$S^{2k+2} \xrightarrow{B^t} S^k \longrightarrow M \longrightarrow 0 \tag{3.2}$$

where B is an admissible matrix.

Any admissible M is a graded S-module and up to twists this grading is determined by the presentation (3.2). Let $e_1, \ldots, e_k$ be the standard basis of S^k and let $u_1, \ldots, u_k$ be the corresponding generators of M. From the matrix B^t it follows that the following relations hold:

$$X_0^{\gamma_{01}} u_1 + X_{l(2)}^{\gamma_{l(2)2}} u_2 = 0,$$
$$X_0^{\gamma_{02}} u_2 + X_{l(3)}^{\gamma_{l(3)3}} u_3 = 0, \ldots, \; X_0^{\gamma_{0k-1}} u_{k-1} + X_{l(k)}^{\gamma_{l(k)k}} u_k = 0.$$

If we twist to give u_1 degree 0, then it follows that to grade M we must have $\deg u_2 = \gamma_{01} - \gamma_{l(2)2}$, $\deg u_3 = \gamma_{01} - \gamma_{l(2)2} + \gamma_{02} - \gamma_{l(3)3}$ or in general,

$$\deg u_i = \sum_{j=1}^{i-1} \gamma_{0j} - \sum_{j=2}^{i} \gamma_{l(j)j}. \tag{3.3}$$

In addition any admissible module M has finite dimension over k since $X_i^p M = 0$ where $p = \sum_{j=1}^{k} \gamma_{ij}$, $0 \leqq i \leqq 3$. We need to prove that an admissible module M admits only one presentation by an admissible matrix. In other words we must show that an admissible module uniquely determines an admissible matrix.

Lemma 3.3. *Let M be an admissible module with a presentation by an admissible matrix B as in* (3.2). *Let $u_1, \ldots, u_k \in M$ be the images of the standard basis $e_1, \ldots, e_k$ of S^k. Suppose $w \in M$ is a homogeneous element such that $w \notin mM$ and $X_3^{\gamma_{31}} w = 0$.* (*Here $m = (X_0, X_1, X_2, X_3)$.*) *Then $w = \alpha u_1$ for some $\alpha \in k$.*

If u is a homogeneous element of M such that $u \notin mM$ and $X_0^{\gamma_{0k}} u = 0$ then $u = \alpha u_k$ for some $\alpha \in k$.

Proof. We will prove this lemma by induction on k. If $k = 1$, then

$$M \cong S/(X_0^{\gamma_{01}}, X_1^{\gamma_{11}}, X_2^{\gamma_{21}}, X_3^{\gamma_{31}}).$$

In this case $M/mM \cong k$ and the condition that w be homogeneous gives that $w = \bar{\alpha}$ for some $\alpha \in k$. Assume $k = 2$. Then first consider $M_1 = M/Su_1$. This module is still admissible, and $\bar{w}$ is a homogeneous element of M_1 such that $X_3^{\gamma_{31}} \bar{w} = 0$. If $\bar{w} = 0$, then $w \in Su_1$. Since w is homogeneous and not in mu_1, it follows that $w = \alpha u_1$ for some $\alpha \in k$. Hence we may assume $\bar{w} \neq 0$. The induction gives that $\bar{w} = \alpha \bar{u}_2$ (since $X_3^{\gamma_{32}} \bar{w} = 0$ as $\gamma_{31} \leqq \gamma_{32}$) where $\alpha \in k$, $\alpha \neq 0$. We may assume $\alpha = 1$ and so $w = au_1 + u_2$ where $a \in S$. We will derive a contradiction from this equation. As $X_3^{\gamma_{31}} u_1 = 0$, $X_3^{\gamma_{31}} w = 0$ we obtain that $X_3^{\gamma_{31}} u_2 = 0$. Hence the relation $(0, X_3^{\gamma_{31}})$ is a combination of the columns of B. Write

$$\begin{aligned}(0, X_3^{\gamma_{31}}) = {} & a_1(X_3^{\gamma_{31}}, 0) + a_2(X_2^{\gamma_{21}}, 0) + a_3(X_0^{\gamma_{01}}, X_{l(1)}^{\gamma_{l(1)2}}) \\ & + a_4(X_1^{\gamma_{11}}, X_{m(1)}^{\gamma_{m(1)2}}) + a_5(0, X_0^{\gamma_{02}}) + a_6(0, X_1^{\gamma_{12}}).\end{aligned}$$

Since $\gamma_{31} \leqq \gamma_{32}$ it follows that either a_3 or $a_4 = 1$. Then as

$$a_1 X_3^{\gamma_{31}} + a_2 X_2^{\gamma_{21}} + a_3 X_0^{\gamma_{01}} + a_4 X_1^{\gamma_{11}} = 0,$$

we reach a contradiction.

Now assume $k > 2$. Again set $M_1 = M/Su_1$. Then $X_3^{\gamma_{32}} \bar{w} = 0$ in M_1 (as $\gamma_{31} \leqq \gamma_{32}$) and so by induction $\bar{w} = \alpha \bar{u}_2$. If $\alpha = 0$, we are done—if not we may assume $\alpha = 1$ and derive a contradiction. If $\alpha = 1$ then $w = au_1 + u_2$ for some $a \in S$. Since $X_3^{\gamma_{31}} w = 0$ and $X_3^{\gamma_{31}} u_1 = 0$, we obtain as before that $X_3^{\gamma_{31}} u_2 = 0$. Then in the module $N = M/(Su_3 + \cdots + Su_k)$, we have $X_3^{\gamma_{31}} \bar{u}_2 = 0$ and by the case $k = 2$ above, we have shown this is impossible. □

The proof of the second claim is exactly the same as the proof above, using the fact that $\gamma_{0i} \geqq \gamma_{0j}$ if $i \leqq j$.

Proposition 3. 4. *Let M be an admissible module and suppose M is presented as in* (3. 2) *by two admissible matrices A and B. Then $A = B$.*

Proof. Clearly both A and B are $k \times (2k+2)$ matrices for the same k. Write A as in Definition 3. 1 ii) and write B as in the same definition but with "′" on the γ_{ij}.

Let $u_1, \ldots, u_k$ be the generating set of M corresponding to the standard basis of S^k in the presentation

$$S^{2k+2} \xrightarrow{A^t} S^k \longrightarrow M \longrightarrow 0,$$

and let $v_1, \ldots, v_k$ be the generating set of M corresponding to the standard basis of S^k in the presentation

$$S^{2k+2} \xrightarrow{B^t} S^k \longrightarrow M \longrightarrow 0$$

of M. We first claim that $u_1 = \alpha v_1$ for some $\alpha \in k$. We know $X_3^{\gamma_{31}} u_1 = 0$ and $X_3^{\gamma'_{31}} v_1 = 0$. Without loss of generality we may assume $\gamma_{31} \leqq \gamma'_{31}$. Then $X_3^{\gamma'_{31}} u_1 = 0$ and u_1 is a homogeneous generator of M. By applying lemma 3. 3 we find that $u_1 = \alpha v_1$ for some $\alpha \in k$. Likewise $u_k = \beta v_k$ for some $\beta \in k$.

We induct on k to prove the proposition. If $k = 1$, then

$$M \cong S/(X_0^{\gamma_{01}}, X_1^{\gamma_{11}}, X_2^{\gamma_{21}}, X_3^{\gamma_{31}}) \cong S/(X_0^{\gamma'_{01}}, X_1^{\gamma'_{11}}, X_2^{\gamma'_{21}}, X_3^{\gamma'_{31}})$$

clearly gives $\gamma_{ij} = \gamma'_{ij}$ and then $A = B$ as we have fixed an ordering of these elements in an admissible matrix.

Next suppose $k = 2$. Then $u_1 = \alpha v_1$, $u_2 = \beta v_2$.

Hence $\gamma_{ij} = \gamma'_{ij}$ and it only remains to see that $l(2) = l'(2)$ and $m(2) = m'(2)$. If not then without loss of generality we may assume

$$A = \begin{pmatrix} X_2^{\gamma_{21}} & X_3^{\gamma_{31}} & X_0^{\gamma_{01}} & X_1^{\gamma_{11}} & 0 & 0 \\ 0 & 0 & X_3^{\gamma_{32}} & X_2^{\gamma_{22}} & X_0^{\gamma_{02}} & X_1^{\gamma_{12}} \end{pmatrix}$$

and

$$B = \begin{pmatrix} X_3^{\gamma_{31}} & X_2^{\gamma_{21}} & X_0^{\gamma_{01}} & X_1^{\gamma_{11}} & 0 & 0 \\ 0 & 0 & X_2^{\gamma_{22}} & X_3^{\gamma_{32}} & X_0^{\gamma_{02}} & X_1^{\gamma_{12}} \end{pmatrix}.$$

Then the column $(X_0^{\gamma_{01}}, X_2^{\gamma_{22}})$ of B is a linear combination of the columns of A with coefficients in S. It is immediate to see this cannot happen.

Now assume $k > 2$. Consider $N' = M/Su_k$. This module is admissible, (since $k > 2$) with presentation given by the matrices $A' =$ first $k-1$ rows of A with the last two columns of zeroes chopped off and $B' =$ first $k-1$ rows of B with the last two columns of zeroes chopped off. This remark follows as $Su_k = Sv_k$ by the above. By induction, $A' = B'$.

Apply the same argument to $N''=M/Su$, N'' is admissible with respect to matrices A'' (resp. B'') which are given by deleting the first row and first two columns of A (resp. B) and possibly switching the terms $X_{l(2)}^{\gamma_{l(2)2}}$, $X_{m(2)}^{\gamma_{m(2)}}$ (resp. $X_{l(2)'}^{\gamma'_{l(2)'2}}$, $X_{m(2)'}^{\gamma'_{m(2)'2}}$) to agree with definition 3.1 ii). In any case the induction shows $A''=B''$ and the last $k-2$ rows of $A=$ last $k-2$ rows of B. From $A'=B'$, the first $k-1$ rows of $A=$ first $k-1$ rows of B. As $k\geqq 3$, this shows $A=B$.

The last proposition we wish to prove in this section is

Proposition 3.5. *Let $C=C(n_1, n_2, n_3)$ be a monomial curve and consider the resolution of $S/I(C)$ as in (1.1). Then the matrix $[\varphi_3]$ given by Proposition 2.8 is admissible. In particular, $\operatorname{Ext}_S^3(S/I(C), S)$ is an admissible module.*

Proof. From the work of [1] and the description above, $[\varphi_3]$ satisfies i), ii), iv) of the definition of admissible matrices. We need only to show that $\gamma_{ij}\geqq\gamma_{ik}$ if $i=0, 1$ and $j\leqq k$ and $\gamma_{ij}\leqq\gamma_{ik}$ if $i=2, 3$ and $j\leqq k$. We will show $\gamma_{ij}\geqq\gamma_{ik}$ if $i=0, 1$ and $j\leqq k$. The other inequality will follow by considering the dual curve $n'_1=n_3-n_2$, $n'_2=n_3-n_1$ and $n'_3=n_3$, with the change of coordinates

$$X_0\longrightarrow X_3,\quad X_1\longrightarrow X_2,\quad X_2\longrightarrow X_1,\quad X_3\longrightarrow X_0.$$

The powers $X_0^{\gamma_{0i}}$ and $X_1^{\gamma_{1i}}$ occur in the matrix $[\varphi_2]$ in the representation of the polynomials in S_{i+2} as the 2×2 minors of a 2×3 matrix of the form

$$A_{i+2}=\begin{pmatrix} a_{11} & a_{12} & -X_1^{\gamma_{1i}} \\ a_{21} & a_{22} & X_0^{\gamma_{0i}} \end{pmatrix}$$

and $f_{i+3}=\Delta_3(A_{i+2})=a_{11}a_{22}-a_{21}a_{12}$. Either $f_{i+2}=(X_0^{\gamma_{0i}}a_{11}-X_1^{\gamma_{1i}}a_{21})$ or

$$f_{\alpha(i+2)}=(a_{12}X_0^{\gamma_{0i}}+a_{22}X_1^{\gamma_{1i}})$$

is in S_{i+3}. Since $X_0^{\gamma_{0i+1}}$ (and $X_1^{\gamma_{1i+1}}$) are determined as the least powers of X_0 (respectively X_1) which occur in any monomial of the three polynomials in S_{i+3} and since $X_0^{\gamma_{0i}}$ and $X_1^{\gamma_{1i}}$ do occur in some binomial in S_{i+3}, it immediately follows that $\gamma_{0i+1}\leqq\gamma_{0i}$ and $\gamma_{1i+1}\leqq\gamma_{1i}$ which is what we needed to show.

4. The Main Theorem

Our purpose in this section is to prove the following theorem.

Theorem 4.1. *Let $C=C(n_1, n_2, n_3)$ be a monomial curve in $\mathbb{P}_k^3$. Set $n=\mu(I(C))$.*

i) *If $S_2\cap S_{n-1}=\emptyset$, then C is the only monomial curve in its even linkage class.*

ii) *The set $S_2\cap S_{n-1}$ contains a binomial of the form $X_0^{\alpha_0}X_3^{\alpha_3}-X_1^{\alpha_1}X_2^{\alpha_2}$ if and only if there are infinitely many distinct monomial curves in the even liaison class of C.*

iii) *If $S_2\cap S_{n-1}\neq\emptyset$ and does not contain a polynomial of the form $X_0^{\alpha_0}X_3^{\alpha_3}-X_1^{\alpha_1}X_2^{\alpha_2}$ then the number of monomial curves in the even liaison class is finite. There can be an arbitrarily large (finite) number of monomial curves in the even liaison class in this case.*

Before we can proceed we need several lemmas. The first lemma is similar to the last proof in the previous section.

Lemma 4. 2. *Write* $[\varphi_2]=(S_{ij})$, *and let* $i\geqq 3$ *be odd. Assume the elements* S_{i-2j}, S_{i-1j}, S_{ij}, *and* S_{i+1j} *of the* j^{th} *column of* $[\varphi_2]$ *are non-zero. Then the* X_0 (*resp.* X_1) *exponent in the terms* $S_{i,j}$ *and* $S_{i+1,j}$ *decrease from the* X_0 (*resp.* X_1) *exponent in the terms* $S_{i-2,j}$ *and* $S_{i-1,j}$. (In fact these exponents decrease by the exponent of the pure power of X_0 (resp. X_1) in the i^{th} or $i+1^{\text{st}}$ row.) *The statement remains valid if one replaces* X_0 *by* X_2 (X_1 *by* X_3) "*decreases*" *by* "*increases*" *and the* i^{th} *or* $(i+1)^{\text{st}}$ *by* $(i-2)^{\text{nd}}$ *or* $(i-1)^{\text{st}}$.

Proof. By [1], given two adjacent 2×3 blocks in $[\varphi_2]$, exactly two columns with nonzero entries are continued from the first to the second block. (Since $|S_i\cap S_{i+1}|=2$, see 2. 3.) Moreover the column with non-zero entries in the second 2×3 block which is preceded by zero entries is the column containing pure powers of X_0 and X_1, while the not continued column with nonzero entries in the first 2×3 block is made up of pure powers of X_2 and X_3. It follows that

$$\begin{vmatrix} X_k^{\gamma_k} & S_{i-2j} \\ X_p^{\gamma_p} & S_{i-1j} \end{vmatrix} = \pm \begin{vmatrix} S_{ij} & -X_1^{\gamma_1} \\ S_{i+1j} & X_0^{\gamma_0} \end{vmatrix}$$

where $\{k, p\}=\{2, 3\}$. Both statements in the lemma follow by comparing forms.

Definition 4. 3. If $C=C(n_1, n_2, n_3)$ is a monomial curve and

$$I(C)=(f_1,\ldots,f_n)$$

is the generating set of $I(C)$ given in 2. 3., then set

$$J_i=J_i(C)=(f_1,\ldots,\hat{f_i},\ldots,f_n:f_i).$$

Lemma 4. 4. *The following statements hold*:

i) $\mu(J_i)=$ *the number of non-zero entries in the* i^{th} *column of* $[\varphi_2]$.

ii) $\sum_{i=1}^{n}\mu(J_i)=6n-12.$

iii) $\mu(J_i)$ *is even for all* i.

iv) $\mu(J_1)=\mu(J_n)=2,\ \mu(J_i)\geqq 4,\ i\neq 1, n.$

v) $\mu(J_i)\geqq 6$ *if and only if*

$$f_i=\pm(X_0^{\gamma_{0t}}X_{m(t+1)}^{\gamma_{m(t+1)t+1}}-X_{l(t+1)}^{\gamma_{l(t+1)t+1}}X_1^{\gamma_{1t}})$$

for some $1\leqq t\leqq n-4$.

Proof. By its definition, J_i is generated by the entries of the i^{th} column of $[\varphi_2]$. Lemma 4. 3 shows that the entries of the i^{th} column minimally generate J_i. This proves i). It follows from i) that

$$\sum_{i=1}^{n}\mu(J_i)=\text{number of nonzero entries in } [\varphi_2].$$

On the other hand each row of $[\varphi_2]$ consists of exactly three elements, those elements being the entries of the 2×3 matrix given in Proposition 2.8. Hence,

$$\sum_{i=1}^{n} \mu(J_i) = 3 \text{ (number of rows of } [\varphi_2])$$
$$= 3\,(2n-4) = 6n-12,$$

which proves ii).

To prove iii) we observe that the non-zero entries in any column occur in parts, being the entries of the columns of 2×3 matrices. This remark proves iii).

To prove iv) we observe that f_1 (respectively f_n) occurs in precisely one of the $S_i (2 \leqq i \leqq n-1)$, namely S_2 (respectively S_{n-1}). Since the rows of $[\varphi_2]$ correspond to relations on the elements of S_i, it follows that there are exactly two relations which involve f_1 (respectively f_n). Therefore $\mu(J_1) = \mu(J_n) = 2$.

On the other hand if $i \neq 1, n$, then $f_i \in S_{i-1} \cap S_i$, and is therefore involved in at least 4 rows of $[\varphi_2]$. Thus $\mu(J_i) \geqq 4$ in this case.

Finally we prove v). First suppose $\mu(J_i) \geqq 6$. Since $\mu(J_i) = 2q$, where $q =$ number of S_j such that $f_i \in S_j$, it follows that $q \geqq 3$ and f_i in at least three S_j. Since $f_i \in S_{i-1}$, S_i, and $f_i \notin S_j$ for $j < i-1$ it must follow that $f_i \in S_{i+1}$, in fact $f_i = f_{\alpha(i+1)}$. If not then $f_i \in S_j$ for some $j > i$. Choose j least with this property. Then as $S_j = \{f_{\alpha(j)} f_j, f_{j+1}\}$, $f_i = f_{\alpha(j)}$. As $|S_{j-1} \cap S_j| = 2$, and $S_{j-1} = \{f_{\alpha(j-1)}, f_{j-1}, f_j\}$, clearly $f_{\alpha(j)} = f_i \in S_{j-1}$, contradicting our least choice. Hence $f_i \in S_{i+1}$, and $f_i = f_{\alpha(i+1)}$. We now apply Proposition 2.8. It follows that

$$f_i = \pm (X_0^{\gamma_{0k}} X_{m(k+1)}^{\gamma_{m(k+1)k+1}} - X_{l(k+1)}^{\gamma_{l(k+1)k+1}} X_1^{\gamma_{1k}})$$

by the construction of $[\varphi_3]$.

Conversely the lemma applied in reverse shows that each

$$X_0^{\gamma_{0k}} X_{m(k+1)}^{\gamma_{m(k+1)k+1}} - X_{l(k+1)}^{\gamma_{l(k+1)k+1}} X_1^{\gamma_{1k}}$$

is $\pm f_i$ for some i, and necessarily $\mu(J_i) \geqq 6$.

Lemma 4.5. *If there is an* i, $1 \leqq i \leqq n$ *with* $\mu(J_i) = 2n-4$, *then* $i = 2$ *or* 3 *and* $f_i \in S_2 \cap S_{n-1}$. *Conversely if* $f_i \in S_2 \cap S_{n-1}$, *then* $\mu(J_i) = 2n-4$.

Proof. First suppose there is an i with $\mu(J_i) = 2n-4$. Since $\mu(J_i) =$ number of elements in the i^{th} column of $[\varphi_2]$ and there are $2n-4$ rows in $[\varphi_2]$, it follows that there is a nonzero entry in every row in the i^{th} column. By the construction of $[\varphi_2]$,

$$f_i \in \bigcap_{j=2}^{n-1} S_j,$$

so that $f_i \in S_2 \cap S_{n-1}$. As $S_2 = \{f_1, f_2, f_3\}$ and $f_1 \notin S_j$ for $j > 2$ we obtain $i = 2$ or 3.

Conversely suppose $f_i \in S_2 \cap S_{n-1}$. The proof of Lemma 4.4 v) shows that if $f_i \in S_j$, $j > i$, then $f_k \in S_n$ for all $i \leqq k \leqq j$. Hence $f_i \in S_2 \cap S_{n-1}$ implies $f_i \in \bigcap_{j=2}^{n-1} S_j$ and so the i^{th} column has a nonzero entry in each row. Thus $\mu(J_i) = 2n-4$.

In the examples 2. 10 and 2. 11, $S_2 \cap S_{n-1} \neq \emptyset$ and so there is an f_i with $\mu(J_i) = 2n-4$. In example 2. 10, $f_2 \in S_2 \cap S_{n-1}$, while in 2. 11, $f_3 \in S_2 \cap S_{n-1}$. If one looks at $[\varphi_2]$ in each example, the column with all non-zero entries is apparent. On the other hand in example 2. 12, $S_2 \cap S_{n-1} = \emptyset$, and a glance at $[\varphi_2]$ shows there is no column with all non-zero entries.

Before we begin the proof of Theorem 4. 1, we will need to discuss how to rebuild the matrix $[\varphi_2]$ when one is given $[\varphi_3]$. The next two lemmas do precisely this.

Lemma 4. 6. *The nonzero entries of the first two rows of* $[\varphi_2]$ *are given by*

$$X_j^{\gamma_{j1}},\ -X_0^{\gamma_{01}} X_i^{\gamma_{i2}-\gamma_{i1}},\ -X_1^t;$$
$$-X_i^{\gamma_{i1}},\ X_1^{\gamma_{11}},\ X_0^{t-s+\gamma_{i1}-\gamma_{j1}} X_j^s$$

where s and t are unknowns and we have written $[\varphi_3]$ *as an admissible matrix in the form of definition* 2. 1. ii).

Proof. By [1] and the description of the resolution given above, the first four × four block of rows and columns of $[\varphi_2]$ is given by

$$\begin{bmatrix} X_j^{\gamma_{j1}} & -S_{12} & -S_{13} & 0 \\ -X_i^{\gamma_{i1}} & S_{22} & S_{23} & 0 \\ 0 & \pm X_i^{\gamma_{i2}} & \pm S_{33} & -X_1^{\gamma_{11}} \\ 0 & \pm X_j^{\gamma_{j2}} & \pm S_{43} & X_0^{\gamma_{01}} \end{bmatrix},$$

where each S_{ij} is a monomial in at most two variables, exactly one of each of S_{12}, S_{22} and S_{13}, S_{23} is a pure power in X_1 (as $f_1 = S_{12}S_{23} - S_{22}S_{13}$) and the distinguished determinant of $[\varphi_3] = X_0^{\gamma_{01}} X_i^{\gamma_{i2}} - X_1^{\gamma_{11}} X_j^{\gamma_{j2}} = S_{12} X_i^{\gamma_{i1}} - S_{22} X_j^{\gamma_{j1}}$. From the above equation, $S_{22} = X_1^{\gamma_{11}}$, $S_{12} = X_0^{\gamma_{01}} X_i^{\gamma_{i2}-\gamma_{i1}}$, and $\gamma_{j2} = \gamma_{j1}$. Since $-S_{12}S_{23} + S_{22}S_{13} = f_1$, clearly S_{13} must be a pure power of X_1, say X_1^t. Then

$$f_1 = X_1^{t+\gamma_{11}} - S_{12}S_{23} = X_1^{t+\gamma_{11}} - X_0^{\gamma_{01}} X_i^{\gamma_{i2}-\gamma_{i1}} S_{23}.$$

On the other hand S_{23} is a monomial which cannot contain a power of X_i or X_1 as f_1 is irreducible and $f_2 = -X_j^{\gamma_{j1}} S_{23} + X_i^{\gamma_{i1}} S_{13}$ is irreducible. Thus $S_{23} = X_0^r X_j^s$ for some s, r, and by homogeneity, $r = t - s + \gamma_{i1} - \gamma_{j1}$, which finishes the proof.

For the next lemma we will assume that $C = C(n_1, n_2, n_3)$ is a monomial curve with $I(C) = (f_1, \ldots, f_n)$ as in Proposition 2. 5 with $n \geqq 5$. Furthermore we will assume $S_2 \cap S_{n-1} \neq \emptyset$. By switching f_2 and f_3 if necessary we may assume $\{f_3\} = S_2 \cap S_{n-1}$ and we will make this assumption throughout the remainder of the paper in the case that $S_2 \cap S_{n-1} \neq \emptyset$.

We wish to simplify notation in this case. Since $\{f_3\} = S_2 \cap S_{n-1}$ it follows as above that $S_i = \{f_3, f_i, f_{i+1}\}$ for $4 \leqq i \leqq n-1$, and the $(2i-3)^{\text{rd}}$ and $(2i-2)^{\text{nd}}$ rows of $[\varphi_2]$ are zero except in the 3^{rd}, i^{th}, and $(i+1)^{\text{st}}$ columns. The nonzero entries in these rows and columns are the matrix A_i, $\Delta_1(A_i) = f_3$, $\Delta_2(A_i) = f_i$, $\Delta_3(A_i) = f_{i+1}$.

Since the distinguished minors of $[\varphi_3]$ are all equal to f_3, we find that $\gamma_{ij}=\gamma_{i1}$ for $i=0, 1$ and $1\leqq j\leqq n-4$ and $\gamma_{ij}=\gamma_{i2}$ for $i=2, 3$ and $2\leqq j\leqq n-3$. Hence $l(i)=l(j)$ for all i, j and $m(i)=m(j)$ for all i, j. We let $l=l(2)$, $m=m(2)$ be the common value. Notice that $f_3=-X_m^{\gamma_m}X_0^{\gamma_0}+X_l^{\gamma_l}X_1^{\gamma_1}$ if we let $\gamma_i=\gamma_{i1}$ for $i=0, 1$, $\gamma_i=\gamma_{i2}$ for $i=2, 3$. The matrix $[\varphi_3]$ now has the form,

$$[\varphi_3]^t=\begin{pmatrix} X_m^{\gamma_m} & 0\dots & 0 & 0 \\ X_l^{\gamma_l} & 0 & & \vdots \\ X_0^{\gamma_0} & X_l^{\gamma_l} & & \\ X_1^{\gamma_1} & X_m^{\gamma_m} & & \\ 0 & X_0^{\gamma_0} & & \\ 0 & X_1^{\gamma_1} & & \\ 0 & 0 & & \\ 0 & 0 & X_l^{\gamma_l} & 0 \\ & & X_m^{\gamma_m} & 0 \\ & & X_0^{\gamma_0} & X_l^{\gamma_l} \\ & & X_1^{\gamma_1} & X_m^{\gamma_m} \\ 0 & & 0 & X_0^{\gamma_{0n-3}} \\ 0\dots & & 0 & X_1^{\gamma_{1n-3}} \end{pmatrix}.$$

Lemma 4.7. *Let $C=C(n_1, n_2, n_3)$ be a monomial curve with $I(C)=(f_1,\dots,f_n)$ as in Proposition 2.5. Furthermore assume $n\geqq 5$ and $S_2\cap S_{n-1}\neq\emptyset$. Write $[\varphi_2]=(S_{ij})$. Then*

$$S_{13}=-X_1^t, \qquad S_{23}=X_l^sX_0^{t-s+\gamma_{m1}-\gamma_l},$$

$$S_{33}=-X_l^{s+\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}-\gamma_0}, \qquad S_{43}=X_m^{\gamma_{m1}}X_1^{t-\gamma_1},$$

$$3\leqq i\leqq n-1 \quad \begin{cases} S_{2i-33}=(-1)^i X_l^{s+(i-2)\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}-(i-2)\gamma_0} \\ S_{2i-23}=(-1)^{i+1}X_m^{\gamma_{m1}+(i-3)\gamma_m}X_1^{t-(i-2)\gamma_1} \end{cases},$$

$$S_{2n-53}=(-1)^{n-1}X_l^{s+(n-3)\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}-(n-4)\gamma_0-\gamma_{0n-3}},$$

$$S_{2n-43}=(-1)^n X_m^{\gamma_{m1}+(n-4)\gamma_m}X_1^{t-(n-4)\gamma_1-\gamma_{1n-3}}$$

gives the entries of the 3rd column of $[\varphi_2]$ for some s and t.

Proof. The form of the first two rows of $[\varphi_2]$ is given in Lemma 4.6 and shows that S_{13} and S_{23} have the required form. Consider the 3rd and 4th rows of $[\varphi_2]$. Since $S_3=\{f_2, f_3, f_4\}$, the only non-zero entries in these rows are

$$A_3=\begin{pmatrix} S_{32} & S_{33} & S_{34} \\ S_{42} & S_{43} & S_{44} \end{pmatrix}.$$

On the other hand, the form of $[\varphi_3]$ together with Proposition 2.8 shows that $S_{34} = -X_1^{\gamma_1}$ and $S_{44} = X_0^{\gamma_0}$. Then

$$\Delta_1(A_3) = f_2 = S_{33}X_0^{\gamma_0} + S_{43}X_1^{\gamma_1} = \Delta_2(A_2) = -X_l^{s+\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}} + X_1^t X_m^{\gamma_{m1}}$$

by Lemma 4.6. Therefore $S_{33} = -X_l^{s+\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}-\gamma_0}$, $S_{43} = X_m^{\gamma_{m1}}X_1^{t-\gamma_1}$ which proves the lemma in the case that $i=3$.

We prove the lemma for $3 \leqq i \leqq n-1$ by induction. First we claim that the non-zero entries (up to sign) in the $(2i-3)^{\text{rd}}$ and $(2i-2)^{\text{nd}}$ rows of $[\varphi_2]$ are

$$\begin{matrix} \begin{pmatrix} S_{2i-33} & -X_m^{\gamma_m} & -X_1^{\gamma_1} \\ S_{2i-23} & X_l^{\gamma_l} & X_0^{\gamma_0} \end{pmatrix} \\ \qquad\quad i^{\text{th}} \quad (i+1)^{\text{st}} \\ \qquad \text{column} \quad \text{column} \end{matrix},$$

where S_{2i-33} and S_{2i-23} are given as in the lemma.

Consider the $(2i-3)^{\text{rd}}$ and $(2i-2)^{\text{nd}}$ rows of $[\varphi_2]$ for $i \geqq 4$. As $S_i = \{f_3, f_i, f_{i+1}\}$ and these rows of $[\varphi_2]$ correspond to the matrix A_i with entries in the 3^{rd}, i^{th}, and $(i+1)^{\text{st}}$ columns of $[\varphi_2]$, we obtain that these rows are of the form,

$$\begin{pmatrix} 0 & 0 & S_{2i-33} & 0\ldots & 0 & S_{2i-3i} & S_{2i-3i+1} & 0\ldots & 0 \\ 0 & 0 & S_{2i-23} & 0\ldots & 0 & S_{2i-2i} & S_{2i-2i+1} & 0\ldots & 0 \end{pmatrix}.$$

Furthermore, since $\Delta_1(A_i) = f_3$ we obtain that

$$S_{2i-3i}S_{2i-2i+1} - S_{2i-2i}S_{2i-3i+1} = f_3.$$

On the other hand, Proposition 2.8 shows that $S_{2i-3i+1} = -X_1^{\gamma_1}$ while $S_{2i-2i+1} = X_0^{\gamma_0}$, for $3 \leqq i \leqq n-2$. We immediately deduce that $S_{2i-3i} = X_l^{\gamma_l}$ and $S_{2i-2i} = -X_m^m$.

Suppose we have shown that S_{2i-33} and S_{2i-23} have the required form for $i \leqq j < n-1$. Then

$$\begin{aligned} f_{j+1} &= \Delta_3(A_j) = S_{2j-33}X_l^{\gamma_l} + S_{2j-23}X_m^{\gamma_m} \\ &= \Delta_2(A_{j+1}) = -S_{2(j+1)-33}X_0^{\gamma_0} - S_{2(j+1)-23}X_1^{\gamma_1}. \end{aligned}$$

As $S_{2j-33} = X_l^{s+(j-2)\gamma_l}X_0^{t-s-\gamma_l+\gamma_{m1}-(j-2)\gamma_0}$ and $S_{2j-23} = -X_m^{\gamma_{m1}+(j-3)\gamma_m}X_1^{t-(j-2)\gamma_1}$, comparing terms in the above equation immediately gives the correct formula for $S_{2(j+1)-33}$ and $S_{2(j+1)-23}$.

This proves the lemma except for the form of S_{2n-53} and S_{2n-43}. We proceed in a similar manner. The last two rows of $[\varphi_2]$ have the form,

$$A_{n-1} = \begin{pmatrix} 0\ldots & 0 & S_{2n-53} & 0\ldots & 0 & S_{2n-5n-1} & S_{2n-5n} \\ 0\ldots & 0 & S_{2n-43} & 0\ldots & 0 & S_{2n-4n-1} & S_{2n-4n} \end{pmatrix}.$$

Again the form of $[\varphi_3]$ shows that $S_{2n-5n} = -X_1^{\gamma_{1n-3}}$ and $S_{2n-4n} = X_0^{\gamma_{0n-3}}$. Since

$$\begin{aligned}\Delta_2(A_{n-1}) &= -S_{2n-53}X_0^{\gamma_{0n-3}} - S_{2n-43}X_1^{\gamma_{1n-3}} = f_{n-1}\\ &= \Delta_3(A_{n-2}) = S_{2n-73}X_l^{\gamma_l} + S_{2n-63}X_m^{\gamma_m},\end{aligned}$$

and comparing coefficients gives the correct formula for S_{2n-53} and S_{2n-43}, and finishes the proof of this lemma.

We now will complete the proof of Theorem 4.1 in case that $n = \mu(I(C)) \geqq 5$. First suppose we are in case i) of the theorem, that is $S_2 \cap S_{n-1} = \emptyset$. Then by Lemma 4.5, $\mu(J_i) < 2n-4$ for all i. By Lemma 4.4 ii),

$$\sum_{i=1}^{n} \mu(J_i) = 6n-12. \tag{4.8}$$

We claim that there exist $i \neq j$ with $\mu(J_i)$, $\mu(J_j) \geqq 6$. If not then suppose $\mu(J_i) < 6$ for all $i \neq j$, with j fixed. Then,

$$\begin{aligned}\sum_{i=1}^{n} \mu(J_i) &= \mu(J_j) + \mu(J_1) + \mu(J_n) + \sum_{\substack{i=2\\ i\neq j}}^{n-1} \mu(J_i)\\ &\leqq \mu(J_j) + 4 + 4(n-3)\end{aligned} \tag{4.9}$$

since $\mu(J_1) = \mu(J_n) = 2$ by Lemma 4.4 iv) and $\mu(J_i)$ is even for all i by Lemma 4.4 iii). Combining (4.8) and (4.9) yields that

$$6n-12 \leqq \mu(J_j) + 4n - 8, \quad \text{or} \quad 2n-4 \leqq \mu(J_j).$$

This inequality contradicts the fact that $S_2 \cap S_{n-1} = \emptyset$. Thus there are $i \neq j$ with $\mu(J_i)$, $\mu(J_j) \geqq 6$.

By Lemma 4.4 v), both f_i and f_j are distinguished minors of the matrix $[\varphi_3]$. Now suppose $C' = C(n_1', n_2', n_3')$ is another monomial curve evenly linked to C. Then $E(C) = E(C')$ up to twists. By Proposition 3.5, $E(C)$ is admissible with admissible matrix $[\varphi_3]$ while $E(C')$ is admissible with matrix $[\varphi_3']$. Thus $[\varphi_3'] = [\varphi_3]$ by Proposition 3.4. In particular f_i, f_j are also distinguished minors of $[\varphi_3']$ and in particular $f_i, f_j \in I(C')$. We will show this implies $C = C'$.

Write

$$f_i = X_0^{p_0}X_l^{p_l} - X_1^{p_1}X_m^{p_m}, \quad f_j = X_0^{q_0}X_a^{q_a} - X_1^{q_1}X_b^{p_b}$$

where $\{l, m\} = \{a, b\} = \{2, 3\}$. Then since f_i and f_j are in $I(C) \cap I(C')$, it follows that

$$p_l n_l = p_1 n_1 + p_m n_m, \quad q_a n_a = q_1 n_1 + p_b n_b \tag{4.10}$$

and

$$p_l n_l' = p_1 n_1' + p_m n_m', \quad q_a n_a' = q_1 n_1' + p_b n_b'. \tag{4.11}$$

Therefore the 3-tuples (n_1, n_2, n_3) and (n_1', n_2', n_3') are both in the kernel of the map $g: \mathbb{Z}^3 \to \mathbb{Z}^2$ given by the matrix

$$g = \begin{pmatrix} p_1 & q_1 \\ \pm p_2 & \pm q_2 \\ \pm p_3 & \pm q_3 \end{pmatrix}$$

where the signs depend upon the equations (4.10) and (4.11).

Either the rank of the submodule $N \subseteqq \mathbb{Z}^3$ generated by (n_1, n_2, n_3) and (n_1', n_2', n_3') is one or $\operatorname{rank}(img) = 1$.

In the latter case the columns of g must be of the form

$$\begin{pmatrix} cr_1 & dr_1 \\ cr_2 & dr_2 \\ cr_3 & dr_3 \end{pmatrix}$$

for some integers r_i, c, d, and c or $d \neq 1$ as $f_i \neq f_j$. Without loss of generality we may assume $c \neq 1$. Then f_i is not irreducible as the equation $X_0^{r_1+r_m-r_l} X_l^{r_l} - X_1^{r_1} X_m^{r_m}$ will divide f_i. This is a contradiction as f_i is a minimal generator of the prime ideal $I(C)$.

Therefore $\operatorname{rank} N = 1$ and so there are positive integers so that

$$c(n_1'', n_2'', n_3'') = (n_1', n_2', n_3') \quad \text{and} \quad d(n_1'', n_2'', n_3'') = (n_1', n_2', n_3').$$

As $(n_1, n_2, n_3) = 1 = (n_1'', n_2'', n_3'')$, it follows that $c = d = 1$ and $n_i = n_i'$. Thus $C = C'$ which proves i).

Now we will prove iii). Suppose $S_2 \cap S_{n-1} = \{-X_0^{\alpha_0} X_2^{\alpha_2} + X_1^{\alpha_1} X_3^{\alpha_3}\}$. We apply Lemma 4.7. First of all, since in the notation of that Lemma,

$$S_2 \cap S_{n-1} = \{-X_m^{\gamma_m} X_0^{\gamma_0} + X_l^{\gamma_l} X_1^{\gamma_1}\}$$

we find that $m = 2$ and $l = 3$. Consider the equation

$$f_n = S_{2n-53} S_{2n-4n-1} - S_{2n-43} S_{2n-5n-1}.$$

This equation must have a pure power of X_2. The formulas for S_{2n-53} and S_{2n-43} in Lemma 4.7 show that with $m = 2$, we must have $t = (n-4)\gamma_1 + \gamma_{1n-3}$ since the exponent of X_1 in S_{2n-43} must be zero to obtain a pure power of X_2 in f_n.

In addition all the exponents in S_{2n-53} are, of course, non-negative so that in particular the exponent of X_0 is non-negative. Therefore

$$t - s - \gamma_l + \gamma_{m1} - (n-4)\gamma_0 - \gamma_{0n-3} \geqq 0.$$

Since $t = (n-4)\gamma_1 + \gamma_{1n-3}$, we obtain that

$$0 \leqq s \leqq (n-4)(\gamma_1 - \gamma_0) + \gamma_{1n-3} + \gamma_{21} - \gamma_3 - \gamma_{0n-3}. \tag{4.12}$$

Thus t is uniquely determined and there are only finitely many choices for s. Also $\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_{1n-3}, \gamma_{21}, \gamma_{0n-3}$ are determined from $[\varphi_3]$ and hence (by Proposition 3.4) from $E = E(C)$. However $\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_{1n-3}, \gamma_{21}, \gamma_{0n-3}, s, t$ determine the matrix $[\varphi_2]$ and hence the defining ideal $I(C)$. If C' is evenly linked to C then $E(C') = E(C)$ implies $\gamma_i' = \gamma_i$ and $\gamma_{1n-3}' = \gamma_{1n-3}$ and therefore by above $t' = t$. Hence the curves evenly linked to C are parametrized by the (finite) choices for s. This proves iii). (We will show there can be arbitrarily large finite families in example 4.19.)

We now will prove ii) of Theorem 4.1. Assume $S_2 \cap S_{n-1} \neq \emptyset$ and contains a binomial of the form $-X_0^{\gamma_0} X_3^{\gamma_3} + X_1^{\gamma_1} X_2^{\gamma_2}$. As above, we know that

$$S_2 \cap S_{n-1} = \{-X_m^{\gamma_m} X_0^{\gamma_0} + X_l^{\gamma_l} X_1^{\gamma_1}\}$$

and so we conclude that $m=3$, $l=2$. As above, since f_n must contain a pure power of X_2, the formulas of Lemma 4.7 show that $\pm S_{2n-53} = X_2^{s+(n-3)\gamma_2}$ and

$$t - s - \gamma_2 + \gamma_{31} - (n-4)\gamma_0 - \gamma_{0n-3} = 0,$$

or

$$s = t - \gamma_2 + \gamma_{31} - (n-4)\gamma_0 - \gamma_{0n-3}. \tag{4.13}$$

In addition we know that $t \geqq (n-4)\gamma_1 + \gamma_{1n-3}$ by considering the exponent of S_{2n-43}.

Consider the first two rows of $[\varphi_2]$. By Lemma 4.6 together with (4.13) we obtain that they are

$$\begin{pmatrix} X_2^{\gamma_2} & -X_0^{\gamma_0} X_3^{\gamma_3 - \gamma_{31}} & -X_1^t \\ -X_3^{\gamma_{31}} & X_1^{\gamma_1} & X_0^{(n-4)\gamma_0 + \gamma_{0n-3}} X_2^{t+c} \end{pmatrix}$$

where $c = \gamma_{31} - \gamma_2 - (n-4)\gamma_0 - \gamma_{0n-3}$.

Therefore the affine relations are,

$$g_1 = X_1^{t+\gamma_1} - X_2^{t+c} X_3^{\gamma_3 - \gamma_{31}}, \quad g_2 = X_2^{t+c+\gamma_2} - X_1^t X_3^{\gamma_{31}}, \quad X_3^{\gamma_3} - X_1^{\gamma_1} X_2^{\gamma_2} = g_3.$$

If $t = -c$ then g_1 and g_2 are the minimal affine relations. If $\gamma_3 = \gamma_{31}$, then g_1 and g_3 are the minimal affine relations.

First suppose g_1 and g_2 are part of a minimal set of affine equations. Then

$$\begin{pmatrix} -(t+\gamma_1) & t+c & \gamma_3 - \gamma_{31} \\ t & -(t+c+\gamma_2) & \gamma_{31} \end{pmatrix}$$

are relations on (n_1, n_2, n_3) and it easily follows as in [2] Lemma 1.2, that the relations above are minimal relations on

$$\begin{aligned} n_1 &= \gamma_{31}(t+c) + (t+c+\gamma_2)(\gamma_3 - \gamma_{31}) \\ &= \gamma_3(t) + \gamma_3 c + \gamma_2 \gamma_3 - \gamma_2 \gamma_{31}, \\ n_2 &= \gamma_{31}(t+\gamma_1) + t(\gamma_3 - \gamma_{31}) \\ &= \gamma_3 t + \gamma_{31}\gamma_1, \\ n_3 &= (t+\gamma_1)(t+c+\gamma_2) - t(t+c) = (\gamma_1 + \gamma_2)t + \gamma_1(c+\gamma_2). \end{aligned} \tag{4.14}$$

From (4.14) it follows that there are constants c_1, c_2, c_3 depending only upon $\gamma_0, \gamma_1, \gamma_3, \gamma_2, c$ such that

$$n_1 = \gamma_3 t + c_1, \quad n_2 = \gamma_3 t + c_2, \quad n_3 = (\gamma_1 + \gamma_2)t + c_3. \tag{4.15}$$

If g_1 and g_3 are minimal affine solutions, a quick check of their defining matrices shows values n_1, n_2, n_3 exactly as in (4.15).

Now consider the equations (4. 15) as *defining* (n_1, n_2, n_3) given t. We know that there is a t_0 such that if n_{01}, n_{02}, n_{03} are the corresponding solutions to (4. 15), then $(n_{01}, n_{02}, n_{03}) = 1$, and $n_{01} < n_{02} < n_{03}$. We make the following three claims:

1) If n_1, n_2, n_3 as in (4. 15) are relatively prime, and $n_1 < n_2 < n_3$, then $C = C(n_1, n_2, n_3)$ is a monomial curve evenly linked to $C_0 = C(n_{01}, n_{02}, n_{03})$.

2) There exist infinitely many values of t such that $(n_{t1}, n_{t2}, n_{t3}) = 1$ where, $n_{t1} = \gamma_3 t + c < n_{t2} = \gamma_3 t + c_2 < n_{t3} = (\gamma_1 + \gamma_2) t + c_3$.

3) The set of $C(n_{t1}, n_{t2}, n_{t3})$ such that $n_{t1} < n_{t2} < n_{t3}$ and $(n_{t1}, n_{t2}, n_{t3}) = 1$ are exactly the set of monomial curves evenly linked to $C = C(n_{01}, n_{02}, n_{03})$.

First we will prove 2). By assumption we have a solution $n_{01} < n_{02} < n_{03}$ with $(n_{01}, n_{02}, n_{03}) = 1$. Let $n_{01} = q_1 d$, $n_{02} = q_2 d$ where $(q_1, q_2) = 1$, $(d, n_{03}) = 1$. Set $t_l = t_0 + ld$, $n_{l_i} = n_{t_l i}$, $1 \leqq i \leqq 3$.

First we claim $n_{l1} < n_{l2} < n_{l3}$. Since $c_1 < c_2$ (as $n_{01} < n_{02}$) $n_{l1} < n_{l2}$ is immediate. On the other hand, $\gamma_3 < \gamma_1 + \gamma_2$ since $\gamma_3 + \gamma_0 = \gamma_1 + \gamma_2$ by homogeneity. Hence clearly $n_{l2} < n_{l3}$.

If $(n_{l1}, n_{l2}) = d = (n_{01}, n_{02})$, then $(n_{l1}, n_{l2}, n_{l3}) = 1$ since

$$(d, n_{l3}) = (d, (\gamma_1 + \gamma_2) t_l + c_3) = (d, (\gamma_1 + \gamma_2) t_0 + c_3) = (d, n_{03}) = 1.$$

On the other hand,

$$\begin{aligned}(n_{l1}, n_{l2}) &= (n_{01} + \gamma_3 ld, n_{02} + \gamma_3 ld) \\ &= d(q_1 + \gamma_3 l, q_2 + \gamma_3 l) \\ &= d(q_1 + \gamma_3 l, q_2 - q_1).\end{aligned}$$

Set $d' = (q_1, \gamma_3)$. For infinitely many l, $q_1 + \gamma_3 l = d' \pi(l)$ where $\pi(l)$ is prime, $\pi(l) > q_2 - q_1$. Thus $(\pi(l), q_2 - q_1) = 1$. Since $d' \mid q_1$, $(d', q_2 - q_1) = 1$ and so for infinitely many l, $(q_1 + \gamma_3 l, q_2 - q_1) = 1$ which proves the second claim.

Next observe that claim 3) follows at once from the work above and claim 1). For if $C = C(n_1', n_2', n_3')$ is evenly linked to $C_0 = C(n_{01}, n_{02}, n_{03})$ then since $E(C_0) = E(C)$ necessarily $\gamma_1, \gamma_2, \gamma_3, \gamma_0, c$ are the same for both curves as $[\varphi_3] = [\varphi_3']$. Then the work above shows that there is some t such that $n_i' = n_{ti}$ as in (4. 14) and (4. 15).

So it remains to prove claim 1). Let n_1, n_2, n_3 be as in (4. 15) for some fixed t, and assume $n_1 < n_2 < n_3$ and $(n_1, n_2, n_3) = 1$.

We now define an ideal J with a resolution where it will transpire that $J = I(C)$. Let $[\varphi_3]$ be the last map in the resolution of $I(C_0)$. We define $[\varphi_2] = (S_{ij})$ as follows: First of all the 3rd column of (S_{ij}) is defined as in Lemma 4. 7 where

$$s = t - \gamma_2 + \gamma_{31} - (n-4)\,\gamma_0 - \gamma_{0n-3}$$

as in (4. 13) and where $m = 3$, $l = 2$. We define $S_{11} = X_2^{\gamma_2}$, $S_{21} = -X_3^{\gamma_{31}}$, $S_{12} = -X_0^{\gamma_0} X_3^{\gamma_3 - \gamma_{31}}$, $S_{22} = X_1^{\gamma_1}$, $S_{32} = X_3^{\gamma_3}$, $S_{42} = -X_2^{\gamma_2}$, $S_{34} = -X_1^{\gamma_1}$, $S_{44} = X_0^{\gamma_0}$ and in general $S_{2i-3i} = -X_3^{\gamma_3}$,

$S_{2i-2i}=X_2^{\gamma_2}$, $S_{2i-3i+1}=-X_1^{\gamma_1}$, $S_{2i-2i+1}=X_0^{\gamma_0}$, (as in Lemma 4.7), and finally $S_{2n-5n-1}=-X_0^{\gamma_3-\gamma_{0n-3}}X_3^{\gamma_3}$, $S_{2n-4n-1}=X_2^{\gamma_2}$, $S_{2n-5n}=-X_1^{\gamma_1}$, and $S_{2n-4n}=X_0^{\gamma_{0n-3}}$ with all other $S_{ij}=0$. The $(2i-3)^{\text{rd}}$ and $(2i-2)^{\text{nd}}$ rows of $[\varphi_2]$ for $i\geqq 4$ are of the form,

$$\begin{pmatrix} 0 & 0 & S_{2i-33} & 0 & \dots & 0 & -X_3^{\gamma_3} & -X_1^{\gamma_1} & 0 & \dots & 0 \\ 0 & 0 & S_{2i-23} & 0 & \dots & 0 & X_2^{\gamma_2} & X_0^{\gamma_0} & 0 & \dots & 0 \end{pmatrix}$$

$$\begin{matrix} & & & & & & \uparrow & \uparrow & & & \\ & & & & & & i^{\text{th}} & (i+1)^{\text{st}} & & & \\ & & & & & & \text{column} & \text{column} & & & \end{matrix}$$

where the 3rd column is given as in Lemma 4.7. Finally we define the ideal $J=$ all the 2×2 binomial minors of the (paired) rows of $[\varphi_2]$; in other words, the signed 2×2 minors from the non-zero entries of the 1st and 2nd rows, 3rd and 4th rows,..., $2n-5$ and $2n-4^{\text{th}}$ rows. We order the elements of J in the obvious fashion and define the map

$$S^n \xrightarrow{[\varphi_1]} S$$

via the n minors we obtain in this fashion. It is easy to check by what was done in [1] (using the Buchsbaum-Eisenbud criterion) that the complex

$$0 \longrightarrow S^{n-3} \xrightarrow{[\varphi_3]} S^{2n-4} \xrightarrow{[\varphi_2]} S^n \xrightarrow{[\varphi_1]} S$$

gives a resolution of S/J. In particular S/J is unmixed. Since $\sqrt{J}=I(C)$ is clear, $I(C)$ is the only associated prime of J. To show $J=I(C)$ it suffices to prove that $J_{I(C)}=I(C)_{I(C)}$. We use a multiplicities argument to show this.

Consider the equations $f_2, f_3 \in J$ which are given by the formulas,

$$f_2=X_1^t X_3^{\gamma_{31}}-X_0^{(n-4)\gamma_0+\gamma_{0n-3}}X_2^{t+\gamma_{31}-(n-4)\gamma_0-\gamma_{0n-3}},$$

$$f_3=X_1^{\gamma_1}X_2^{\gamma_2}-X_0^{\gamma_0}X_3^{\gamma_3}.$$

We consider the residual intersection of $I(C)$ with (f_2, f_3). We find

$$\sqrt{(f_2, f_3)}=I(C)\cap(X_1, X_0)\cap(X_2, X_3)$$

and so

$$\begin{aligned}(4.16)\qquad (\deg f_2)(\deg f_3) &= e(S/(f_2, f_3)) \\ &= e(S/I(C))\, l((S/(f_2, f_3))_{I(C)}) \\ &\quad + e(S/(X_0, X_1))\, l((S/(f_2, f_3))_{(X_0, X_1)}) \\ &\quad + e(S/(X_2, X_3))\, l((S/(f_2, f_3))_{(X_2, X_3)})\end{aligned}$$

where $e(\ )$ denotes the multiplicity and $l(\ )$ denotes length.

By the remark (4.26) below we find that

$$l((S/(f_2, f_3))_{(X_0, X_1)})=\min(((n-4)\gamma_0+\gamma_{0n-3})\gamma_1, \gamma_0 t)$$

while

$$l((S/(f_2, f_3))_{(X_2, X_3)})=\min(\gamma_{31}\gamma_2, (t+\gamma_{31}-(n-4)\gamma_0-\gamma_{0n-3})\gamma_3).$$

We claim that $((n-4)\gamma_0+\gamma_{0n-3})\gamma_1 \leqq \gamma_0 t$. This follows as $n_3 > n_2$ implies that

$$(\gamma_1+\gamma_2)t+\gamma_1(\gamma_{31}-(n-4)\gamma_0-\gamma_{0n-3}) > \gamma_3 t+\gamma_{31}\gamma_1$$

or

$$(\gamma_1+\gamma_2-\gamma_3)t > \gamma_1((n-4)\gamma_0+\gamma_{0n-3})$$

which is the required inequality since $\gamma_1+\gamma_2-\gamma_3=\gamma_0$ by the homogeneity of C_0. Also we observe that

$$\gamma_{31}\gamma_2 \leqq (t+\gamma_{31}-(n-4)\gamma_0-\gamma_{0n-3})\gamma_3$$

since $\gamma_3 \geqq \gamma_{31}$ holds and $t \geqq (n-4)\gamma_0+\gamma_{0n-3}-\gamma_{31}-\gamma_2$. Hence, (4. 16) reads,

$$\begin{aligned}(t+\gamma_{31})(\gamma_1+\gamma_2) &= (\deg f_2)(\deg f_3)\\ &= n_3 l(S/(f_2, f_3)_{I(C)})\\ &\quad +((n-4)\gamma_0+\gamma_{0n-3})\gamma_1+\gamma_{31}\gamma_2,\end{aligned}$$

or,

$$n_3 l(S/(f_2, f_3)_{I(C)}) = (\gamma_1+\gamma_2)t+\gamma_1(\gamma_{31}-(n-4)\gamma_0-\gamma_{0n-3}).$$

Since the right-hand side of the above equation is equal to n_3 (see 4. 14) it follows that

$$l(S/(f_2, f_3)_{I(C)}) = 1.$$

Therefore $l((S/J)_{I(C)})=1$ also as $(f_2, f_3) \subseteq J$ and so $J=I(C)$. Hence $E(C)=\operatorname{coker}[\varphi_3]^t=E(C_0)$ and C and C_0 are evenly linked. This concludes the proof of Theorem 4. 1 iii) in the case that $n \geqq 5$. In fact claims 1) and 2) establish the Theorem, but part 3) shows that the formulas (4. 15) give precisely the set of monomial curves evenly linked to $C=C_0$.

Before we finish the proof of Theorem 4. 1, we will give several examples to illustrate what we have proved so far.

First we return to the examples 2. 10, 2. 11.

Example 2. 10. We let $C=C(2, 5, 7)$. This curve satisfies iii) of Theorem 4. 1 as $-X_0X_3+X_1X_2 \in \cap S_i$. In the notation of the proof of this theorem, $\gamma_0=\gamma_1=\gamma_2=\gamma_3=\gamma_{0n-3}=\gamma_{31}=1$ and $n=5$. Hence all the monomial curves evenly linked to $C(2, 5, 7)$ are parametrized by the equations (4. 14) which in this case become,

$$n_1=t-2, \quad n_2=t+1, \quad n_3=2t-1, \tag{4. 17}$$

and we have the further conditions that $1 \leqq n_1 < n_2 < n_3$ and $(n_1, n_2, n_3)=1$. If we let $a=t-2$, then this set of curves is exactly the set

$$\{C(a, a+3, 2a+3) \mid a \geqq 1,\ (a, 3)=1\}.$$

The curve $C(2, 5, 7)$ is the case $a=2$. As one can also show that $C(2, 5, 7)$ has no monomial curves in its odd linkage class, the above set gives all the monomial curves in the liaison class of $C(2, 5, 7)$.

Example 2. 11. Here $C=C(1, 5, 8)$ satisfies ii) of Theorem 4. 1 as

$$S_2 \cap S_{n-1} = \{X_1^2 X_3 - X_0 X_2^2\}.$$

In this case, $\gamma_0 = \gamma_3 = 1$, $\gamma_1 = \gamma_2 = 2$, $\gamma_{1n-3} = \gamma_{12}$ (as $n=5$) $= 1$ and $\gamma_{21} = 1$.

As in the proof of Theorem 4. 1, we find that $t=3$ and $0 \leqq s \leqq 1$ (see 4. 12). If $s=0$ we obtain the curve $C(1, 5, 8)$ while if $s=1$ we obtain the curve $C(3, 7, 8)$. Therefore the entire monomial liaison class consists of $C(1, 5, 8)$ and $C(3, 7, 8)$.

Example 4. 18. In this example we give examples of infinite monomial liaison classes with any number of generators of $I(C)$.

Let $E(p)$ be the admissible module presented by the matrix $[\varphi_3]^t$:

$$S^{2p+2}(-1) \xrightarrow[[\varphi_3]^t]{} S^p \longrightarrow E(p) \longrightarrow 0$$

where

$$[\varphi_3] = \begin{pmatrix} X_2 & X_3 & X_0 & X_1 & 0 & 0\ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & X_2 & X_3 & X_0 & X_1 & 0\ldots & & & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & 0 & \ldots & & & & X_2 & X_3 & X_0 & X_1 \end{pmatrix}.$$

The reconstruction of Theorem 4. 1 shows that $E = E(C(1, p+2, p+3))$. To find the liaison class of $C(1, p+2, p+3)$ we note that

$$\gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_{0p} = \gamma_{31} = 1,$$

$n = p+3$. Therefore this liaison class is parametrized as in (4. 14) by

$$n_1 = t-p, \quad n_2 = t+1, \quad n_3 = 2t+1-p.$$

If we set $a = t-p$, then this class is

$$\{C(a, a+p+1, 2a+p+1) \mid a \geqq 1,\ (a, p+1)=1\}.$$

Example 4. 19. In this example we show that there may be liaison classes of monomial curves which are finite but have an arbitrarily high number of monomial curves in the liaison class.

Let p, q be two integers such that

$$1 \leqq q < 3p-9, \quad (q, 3p)=1, \quad (q+6, p-1)=1.$$

Define $q(p-1)=n_1$, $p(q+6)=n_2$, $3p(p-1)=n_3$. We first observe that $n_1 < n_2 < n_3$ and $(n_1, n_2, n_3)=1$ both follow from our assumptions.

Consider the curves $C_{p,q} = C(n_1, n_2, n_3)$ where p is fixed and q is allowed to vary subject to the conditions above.

We claim that $I(C_{p,q})=(f_1,\dots,f_5)$ where

$$f_1=X_1^{3p}-X_0^{3p-q}X_3^q, \qquad f_2=-X_0^{3p-q-3}X_3^{q+2}+X_2^{p-1}X_1^{2p},$$
$$f_3=-X_0^3X_2^{p-1}+X_1^pX_3^2, \qquad f_4=X_1^pX_2^{2p-2}-X_0^{3p-q-6}X_3^{q+4},$$
$$f_5=-X_2^{3p-3}+X_0^{3p-q-9}X_3^{q+6},$$

and we claim the resolution of $J=(f_1,\dots,f_5)$ is

(4. 20)

$$0 \longrightarrow \bigoplus_{S(-4p)} S(-4p-1) \xrightarrow{[\varphi_3]} \begin{matrix} S(-4p+1)\\ \oplus\\ S(-3p-2)\\ \oplus\\ S(-3p-1)\\ \oplus\\ S(-4p+2)\\ \oplus\\ S(-3p)\\ \oplus\\ S(-4p+3)\end{matrix} \xrightarrow{[\varphi_2]} \begin{matrix} S(-3p)\\ \oplus\\ S(-3p+1)\\ \oplus\\ S(-p-2)\\ \oplus\\ S(-3p+2)\\ \oplus\\ S(-3p+3)\end{matrix} \xrightarrow{[\varphi_1]} S$$

where $[\varphi_1]=(f_1,\dots,f_5)^t$,

$$[\varphi_2]=\begin{pmatrix} X_3^2 & -X_0^3 & -X_1^{2p} & 0 & 0\\ -X_2^{p-1} & X_1^p & X_0^{3p-q-3}X_3^q & 0 & 0\\ 0 & X_2^{p-1} & -X_0^{3p-q-6}X_3^{q+2} & -X_1^p & 0\\ 0 & -X_3^2 & X_2^{p-1}X_1^p & X_0^3 & 0\\ 0 & 0 & X_0^{3p-q-9}X_3^{q+4} & -X_2^{p-1} & -X_1^p\\ 0 & 0 & -X_2^{2p-2} & X_3^2 & X_0^3 \end{pmatrix}.$$

and

$$[\varphi_3]=\begin{pmatrix} X_2^{p-1} & X_3^2 & X_0^3 & X_1^p & 0 & 0\\ 0 & 0 & X_3^2 & X_2^{p-1} & X_0^3 & X_1^p \end{pmatrix}.$$

We leave it to the reader to check that (4. 20) is indeed a resolution of S/J. (It is easy using the Buchsbaum-Eisenbud criterion.) It is also trivial to check that $\sqrt{J}=I(C_{p,q})$. To show $J=I(C_{p,q})$ we use a multiplicities argument.

Consider the residual intersection of J with (f_1,f_3). We find that

$$\sqrt{(f_1,f_3)}=I(C_{p,q})\cap(X_0,X_1)$$

and by Remark 4. 26, $l((S/(f_1,f_3))_{(X_0,X_1)})=\min\{9p,\ p(3p-q)\}=9p$ (as $q<3p-9$). Hence,

$$\begin{aligned} e(S/(f_1,f_3)) &= \deg(f_1)\deg(f_3)\\ &= l((S/(f_1,f_3))_{I(C_{p,q})})\cdot e(S/I(C_{p,q}))\\ &\quad + l((S/(f_1,f_3))_{(X_0,X_1)})\cdot e(S/(X_0,X_1)). \end{aligned}$$

Consequently,

$$3p(p+2) = l((S/(f_1, f_3))_{I(C_{p,q})}) \cdot 3p(p-1) + 9p$$

since $3p(p-1) = n_3 = e(S/J(C_{p,q}))$.

It follows that

$$l((S/(f_1, f_3))_{I(C_{p,q})}) = 1$$

and so in particular $J_{I(C_{p,q})} = (f_1, f_3)_{I(C_{p,q})} = I(C_{p,q})$. The resolution (4. 20) of S/J shows S/J is unmixed. As $\sqrt{J} = I(C_{p,q})$, it follows that $J = I(C_{p,q})$ which we claimed.

Now consider $E(C_{p,q})$ and the liaison class of $C_{p,q}$. Since

$$S_2 \cap S_4 = \{f_3\} = \{X_3^2 X_1^p - X_0^3 X_2^{p-1}\}$$

Theorem 4. 1 iii) gives that the even liaison class of $E(C_{p,q})$ is finite. On the other hand, the matrix $[\varphi_3]$ does not depend upon q, so that for fixed p, $E(C_{p,q'}) = E(C_{p,q})$ (even as graded modules without twists). Thus for fixed p, there are at least as many monomial curves in the even liaison class as there are values of q which satisfy $1 \leqq q < 3p - q$, $(q, 3p) = 1$ and $(q+6, p-1) = 1$. For instance if $p = 10$, then the conditions are $1 \leqq q < 21$ $(q, 30) = 1$, $(q+6, 9) = 1$ we find the following values for q: $q = 1, 7, 11, 13, 17, 19$. The corresponding monomial primes are

$$C_{10,1} = C(9, 70, 270), \qquad C_{10,7} = C(63, 130, 270),$$
$$C_{10,11} = C(99, 170, 270), \qquad C_{10,13} = C(117, 190, 270),$$
$$C_{16,17} = C(153, 230, 270), \qquad C_{10,19} = C(171, 250, 270).$$

These form exactly one liaison class of monomial curves. (As indicated before it follows that whenever $\mu(I(C)) = 5$, there are no monomial curves in the odd liaison class of C.)

It is also easy to see that the liaison class can be made arbitrarily large by taking $q = 5^m - 6$, $p = 5^n$, $2 \leqq m < n$.

We will finish the proof of Theorem 4. 1 by proving the following theorem.

Theorem 4. 21. *Suppose $C = C(n_1, n_2, n_3)$ is a monomial curve and $E = E(C)$, so that $E \cong S/(X_0^{\gamma_0}, X_1^{\gamma_1}, X_2^{\gamma_2}, X_3^{\gamma_3})$ (as E is admissible). Then,*

i) $\gamma_1 + \gamma_2 \geqq \gamma_0 + \gamma_3$.

ii) *There are integers p, q such that*

a) $p \geqq \max\{0, \gamma_1 + \gamma_3 - \gamma_0 - \gamma_2, \gamma + 1\}$ *where*

$$\gamma = \frac{\gamma_3(\gamma_1 + \gamma_2) + \gamma_2(q - \gamma_0 - \gamma_1 - \gamma_2) + q(\gamma_0 - \gamma_3)}{(\gamma_1 + \gamma_2 - q)}.$$

b) $\gamma_3 \leqq q \leqq \gamma_1 + \gamma_2 - \gamma_0$.

c) $(\gamma_1, \gamma_2, q) = 1$, $(p+\gamma_2, \gamma_0, \gamma_3) = 1$,

$(p, q-\gamma_3, \gamma_0+\gamma_1+\gamma_2-\gamma_3) = 1$,

$(p+2\gamma_2, q+\gamma_3, \gamma_0-\gamma_1-\gamma_2-\gamma_3) = 1$.

d) $n_1 = pq+\gamma_2(q-\gamma_3)$,

$n_2 = n_1+\gamma_3(\gamma_1+\gamma_2)+q(\gamma_0-\gamma_3)$,

$n_3 = (\gamma_1+\gamma_2)\,p+\gamma_2(\gamma_0+\gamma_1+\gamma_2-\gamma_3)$.

Furthermore suppose p' and q' are integers satisfying ii) a)—c) *and we define n_1', n_2', n_3' as in* d) *using p' and q'. If $(n_1', n_2', n_3') = 1$, then C is evenly linked to $C' = C(n_1', n_2', n_3')$. In particular C is evenly linked to infinitely many distinct monomial curves.*

Proof. By Lemma 4.6, the first two rows of $[\varphi_2]$, are of the form

$$\begin{pmatrix} X_3^{\gamma_3} & -X_0^{\delta}X_2^{p} & -X_1^{b} & 0 \\ -X_2^{\gamma_2} & X_1^{a} & X_0^{c}X_3^{d} & 0 \end{pmatrix}.$$

From this and the fact that

$$[\varphi_3] = (X_2^{\gamma_2} \quad X_3^{\gamma_3} \quad X_0^{\gamma_0} \quad X_1^{\gamma_1})$$

we deduce that $[\varphi_2]$ has the form

$$[\varphi_2] = \begin{pmatrix} X_3^{\gamma_3} & -X_0^{\delta}X_2^{p} & -X_1^{b} & 0 \\ -X_2^{\gamma_2} & X_1^{a} & X_0^{c}X_3^{d} & 0 \\ 0 & X_0^{\delta-\gamma_0}X_2^{p+\gamma_2} & -X_0^{c-\gamma_0}X_3^{d+\gamma_3} & -X_1^{\gamma_1} \\ 0 & -X_1^{a-\gamma_1}X_3^{\gamma_3} & X_1^{b-\gamma_1}X_2^{\gamma_2} & X_0^{\gamma_0} \end{pmatrix}.$$

Since the 2×2 minor, $\Delta_3(A_2)$, of the last two rows must have a pure power of X_2, we deduce that $b = \gamma_1$ and $\delta = \gamma_0$. Also all the exponents must be nonnegative so that $a \geqq \gamma_1$, $c \geqq \gamma_0$. By homogeneity,

$$a+\gamma_3 = \gamma_2+\delta+p = \gamma_2+\gamma_0+p$$

so that we obtain

$$\gamma_2+\gamma_0+p-\gamma_3 \geqq \gamma_1, \quad c \geqq \gamma_0. \tag{4.22}$$

The only undetermined coefficients are p and d, as $a = \gamma_2+\gamma_0+p-\gamma_3$ and

$$\begin{aligned} c &= a+b-\delta-p-d \\ &= (\gamma_2+\gamma_0+p-\gamma_3)+\gamma_1-\gamma_0-p-d \\ &= \gamma_2+\gamma_1-\gamma_3-d. \end{aligned}$$

Set $q = d+\gamma_3$. To summarize we have,

$$(*) \qquad \begin{aligned} &a = \gamma_2+\gamma_0+p-\gamma_3, \quad b = \gamma_1, \\ &c = \gamma_2+\gamma_1-q, \quad d = q-\gamma_3, \quad \delta = \gamma_0. \end{aligned}$$

The inequalities (4.22) show that $\gamma_2+\gamma_1-q \geqq \gamma_0$ and as $q-\gamma_3 = d \geqq 0$, it follows that $\gamma_2+\gamma_1 \geqq \gamma_0+\gamma_3$ which is condition i) of the theorem. These two inequalities also give

ii) b). As $a \geqq \gamma_1$, we obtain from (*) that $\gamma_2+\gamma_0+p-\gamma_3 \geqq \gamma_1$ or $p \geqq \gamma_3+\gamma_1-\gamma_2-\gamma_0$. Thus ii) a) is satisfied except to show that $p \geqq \gamma+1$.

From the first two rows of $[\varphi_2]$, we obtain that

$$f_2 = X_1^{\gamma_1} X_2^{\gamma_2} - X_0^{\gamma_1+\gamma_2-q} X_3^q,$$
$$f_3 = X_1^{\gamma_0+\gamma_2-\gamma_3+p} X_3^{\gamma_3} - X_0^{\gamma_0} X_2^{\gamma_2}$$

are minimal generators of $I(C)$. It follows that n_1, n_2, n_3 are determined as the 2×2 minors of the matrix of exponents, (for example see [2], Lemma 1.3 and the implied algorithm as explained on p. 6 of [2])

$$\begin{pmatrix} \gamma_1 & \gamma_2 & -q \\ -(\gamma_0+\gamma_2-\gamma_3+p) & p+\gamma_2 & -\gamma_3 \end{pmatrix}.$$

Therefore,

$$n_1 = pq+\gamma_2(q-\gamma_3),$$
$$n_2 = n_1+\gamma_3(\gamma_1+\gamma_2)+q(\gamma_0-\gamma_3),$$
$$n_3 = (\gamma_1+\gamma_2)\,p+\gamma_2(\gamma_0+\gamma_1+\gamma_2-\gamma_3).$$

This proves iii). The relations $n_3 > n_2$ and $q \leqq \gamma_1+\gamma_2-\gamma_0$ show that

$$p > \frac{[\gamma_3(\gamma_1+\gamma_2)+\gamma_2(q-\gamma_0-\gamma_1-\gamma_2)+q(\gamma_0-\gamma_3)]}{(\gamma_1+\gamma_2-q)}$$

which finishes the proof of ii) a).

It remains to prove ii) c). However we know that $(n_1, n_2, n_3) = 1$. We leave it to the reader to verify that this condition translates directly into ii) c).

Now suppose p', q' are integers satisfying ii) a)—c) and we define n_i' as in ii) d) using p' and q' in place of p and q. We need to show that $C' = C(n_1', n_2', n_3')$ is linked to C.

First of all the conditions in ii) a)—c) force $n_1' < n_2' < n_3'$ and we are assuming that $(n_1', n_2', n_3') = 1$. Consider the matrices

$$[\varphi_2'] = \begin{pmatrix} X_3^{\gamma_3} & -X_2^{p'} X_0^{\gamma_0} & -X_1^{\gamma_1} & 0 \\ -X_2^{\gamma_2} & X_1^{\gamma_2+\gamma_0-\gamma_3+p'} & X_0^{\gamma_2+\gamma_1-q'} X_3^{q'} & 0 \\ 0 & X_2^{p'+\gamma_2} & -X_3^{q'} X_0^{\gamma_2+\gamma_1-\gamma_0-q'} & -X_1^{\gamma_1} \\ 0 & -X_3^{\gamma_3} X_1^{\gamma_2+\gamma_0-\gamma_1-\gamma_3+p'} & X_2^{\gamma_2} & X_0^{\gamma_0} \end{pmatrix}$$

and $[\varphi_1']^t = (f_1', f_2', f_3', f_4')$ where

$$f_1' = X_1^{\gamma_1+\gamma_2+\gamma_0-\gamma_3+p'} - X_0^{\gamma_2+\gamma_1+\gamma_0-q} X_2^{p'} X_3^{q'-\gamma_3},$$
$$f_2' = X_0^{\gamma_2+\gamma_1-q'} X_3^{q'} - X_1^{\gamma_1} X_2^{\gamma_2},$$
$$f_3' = X_1^{\gamma_2+\gamma_0-\gamma_3+p'} X_3^{\gamma_3} - X_2^{p'+\gamma_2} X_0^{\gamma_0},$$
$$f_4' = X_2^{p'+2\gamma_2} - X_0^{\gamma_2+\gamma_1-\gamma_0-q'} X_1^{\gamma_2+\gamma_0-\gamma_1-\gamma_3+p'} X_3^{q'+\gamma_3},$$
$$[\varphi_3'] = [\varphi_3] = (X_2^{\gamma_2} X_3^{\gamma_3} X_0^{\gamma_0} X_1^{\gamma_1}).$$

A simple check (using the criterion of Buchsbaum and Eisenbud) shows that

$$(4.23) \qquad 0 \longrightarrow S \xrightarrow{[\varphi'_3]} S^4 \xrightarrow{[\varphi'_2]} S^4 \xrightarrow{[\varphi'_1]} S \longrightarrow S/J \longrightarrow 0$$

is a resolution of S/J where $J=(f'_1, f'_2, f'_3, f'_4)$. Furthermore the way n'_i are defined as in iii) shows that $J \subseteqq I(n'_1, n'_2, n'_3)$. We only need to show that

$$J = I(n'_1, n'_2, n'_3)$$

since in this case $E(C') \cong S/\mathrm{im}((\varphi'_3)^t) = S/(X_0^{\gamma_0}, X_1^{\gamma_1}, X_2^{\gamma_2}, X_3^{\gamma_3}) \cong E(C)$. (All up to twists.)

That $J = I(n'_1, n'_2, n'_3)$ follows from Lemma 1.2 of [2] since the n'_i are the relatively prime 2×2 minors of the matrix of exponents giving f'_2 and f'_3.

However we will give another argument. The resolution (4.23) shows that every associated prime of J has height 2. An elementary calculation shows that

$$\sqrt{J} = I = I(n'_1, n'_2, n'_3)$$

so that I is the only associated prime of J. Consider the equations (f'_2, f'_3). These form a regular sequence and if we denote the multiplicity by $e(\ )$,

$$(4.24) \qquad e(S/(f'_2, f'_3)) = (\deg f'_2)\,(\deg f'_3) = (\gamma_1+\gamma_2)\,(p'+\gamma_2+\gamma_0).$$

On the other hand $\sqrt{(f'_2, f'_3)} = I \cap (X_0, X_1) \cap (X_2, X_3)$. Hence

$$(4.25) \qquad e(S/(f'_2, f'_3)) = l(S_I/(f'_2, f'_3)_I)\cdot e(S/I) + l((S/(f'_2, f'_3))_{(X_0, X_1)}) + l((S/(f'_2, f'_3))_{(X_2, X_3)})$$

since $e(S/(X_0, X_1)) = e(S/(X_2, X_3)) = 1$. We use the following remark whose proof we leave to the reader.

Remark 4.26. Let T be a regular 2-dimensional local ring whose maximal ideal is generated by X, Y. Suppose $f = \alpha_1 X^{k_1} + \alpha_2 Y^{k_2}$ and $g = \beta_1 X^{l_1} + \beta_2 Y^{l_2}$ are a system of parameters in T where $l_i, k_i \geqq 1$, α_i, β_i are units. Then $l(T/(f, g)) = \min\{k_1 l_2, k_2 l_1\}$.

Using this remark, we find that

$$l((S/(f'_2, f'_3))_{(X_0, X_1)}) = \min\{\gamma_0\gamma_1, (\gamma_2+\gamma_1-q')\,(\gamma_2+\gamma_0-\gamma_3+p')\} = \gamma_0\gamma_1$$

since $\gamma_2+\gamma_1-q' \geqq \gamma_0$ by ii) b) and $\gamma_2+\gamma_0-\gamma_3+p' \geqq \gamma_1$ by ii) a).

Also, $l((S/(f'_2, f'_3))_{(X_2, X_3)}) = \min\{\gamma_2\gamma_3, (p'+\gamma_2)\,q'\} = \gamma_2\gamma_3$ since $q' \geqq \gamma_3$ by ii) b).

Combining these remarks we find that

$$\begin{aligned} e(S/I)\, l(S_I/(f'_2, f'_3)_I) &= (\gamma_1+\gamma_2)\,(p'+\gamma_2+\gamma_0) - \gamma_2\gamma_3 - \gamma_1\gamma_0 \\ &= (\gamma_1+\gamma_2)\,p' + \gamma_2(\gamma_0+\gamma_1+\gamma_2-\gamma_3) = n'_3 \quad \text{by ii) d).} \end{aligned}$$

Since $e(S/I) = n'_3$ it follows that $l(S_I/(f'_2, f'_3)_I) = 1$, and thus $l(S_I/J_I) = 1$ also. It immediately follows that $J = I$.

Example 4.27. The module $E = S/(X_0^2, X_1, X_2, X_3)$ is not the Hartshorne-Rao module of any monomial curve since $\gamma_1+\gamma_2 = 2 \ngeqq \gamma_0+\gamma_3 = 3$ contradicting Theorem 4.21 i).

Example 4. 28. Let us find the monomial curves evenly linked to $C = C(1, 4, 6)$. The resolution of $S/I(C)$ shows that

$$E(C) \cong S/(X_0, X_1^2, X_2, X_3).$$

To find the general monomial curve with this Rao module we apply the Theorem above. We find $\gamma_3 = 1 \leqq q' \leqq 2 = \gamma_1 + \gamma_2 - \gamma_0$. If $q' = 1$, then the condition ii) gives $p' \geqq \max\{0, 1, 1\}$, $(p' + 1, 1, 1) = 1$, $(p', 0, 3) = 1$ and $(p', 2, -3) = 1$ which combine to give $p' \geqq 1$, $(p', 3) = 1$. In this case we find $n_1' = p'$, $n_2' = p' + 3$, $n_3' = 3p' + 3$.

If $q' = 2$, then $p' \geqq \max\{0, 1, 2\}$, $(p', 1, 3) = 1$, $(p' + 2, 3, -3) = 1$; thus $p' \geqq 2$ and $p' \not\equiv 1 \bmod 3$. In this case $n_1' = 2p' + 1$, $n_2' = 2p' + 4$, $n_3' = 3p + 3$. Thus the set of monomial curves in the liaison class of $C(1, 4, 6)$ is exactly the set

$$\{C(p, p+3, 3p+3) \mid p \geqq 1,\ (p, 3) = 1\}$$
$$\cup \{C(2p+1, 2p+4, 3p+3) \mid p \geqq 2,\ p \not\equiv 1 \bmod 3\}.$$

Example 4. 29. If $C = C(1, 3, 4)$, then $E(C) \cong S/(X_0, X_1, X_2, X_3) \cong k$. The conditions of Theorem 4. 21 show that the set of monomial curves in the liaison class of C is $\{C(p, p+2, 2p+2) \mid p > 0,\ (p, 2) = 1\}$. This class is exactly the class of arithmetically Buchsbaum curves.

Appendix

In this appendix we summarize the results of this paper to provide an easy algorithm for determining the entire even monomial liaison class of a monomial curve $C = C(n_1, n_2, n_3)$. There are five cases we must separate:

a) $\mu(I(C)) = 3$,

b) $\mu(I(C)) = 4$,

c) $\mu(I(C)) \geqq 5$ and $S_2 \cap S_{n-1} = \emptyset$,

d) $\mu(I(C)) \geqq 5$ and $S_2 \cap S_{n-1} \neq \emptyset$ with $S_2 \cap S_{n-1} = \{X_0^{\alpha_0} X_3^{\alpha_3} - X_1^{\alpha_1} X_2^{\alpha_2}\}$,

e) $\mu(I(C)) \geqq 5$ and $S_2 \cap S_{n-1} \neq \emptyset$ with $S_2 \cap S_{n-1} = \{X_0^{\alpha_0} X_2^{\alpha_2} - X_1^{\alpha_1} X_3^{\alpha_3}\}$.

To determine which case is applicable is a simple matter. First compute the sets S_2 (with corresponding matrix A_2) and S_{n-1} outlined in Section 2; these sets are determined by the affine pieces of C, $X_0 = 1$ and $X_3 = 1$ respectively. If $S_2 = S_{n-1}$ then we are in case a) and the liaison class of C consists of all arithmetically Cohen-Macaulay curves. If $|S_2 \cap S_{n-1}| = 2$, we are in case b), while if $|S_2 \cap S_{n-1}| = 1$ we are in either case d) or e) and if $S_2 \cap S_{n-1} = \emptyset$ we are clearly in case c). In case c) the curve C is the *only* monomial curve in its even liaison class. We now treat b), d), and e) separately.

Case b). In this case the matrix A_2 has the form,

$$\text{(A. 1)} \qquad \begin{pmatrix} X_3^{\gamma_3} & -X_0^{\gamma_0} X_2^{p} & -X_1^{\gamma_1} \\ -X_2^{\gamma_2} & X_1^{\gamma_2 + \gamma_0 + p - \gamma_3} & X_0^{\gamma_2 + \gamma_1 - q} X_3^{q - \gamma_3} \end{pmatrix}.$$

This allows us to solve for $\gamma_0, \gamma_1, \gamma_2, \gamma_3$. (See Theorem 4.21.) Also we may solve for p, q. Then the even liaison class of C is the set of all $C(m_1, m_2, m_3)$ where

$$\begin{aligned} m_1 &= ab + \gamma_2(b - \gamma_3), \\ m_2 &= m_1 + \gamma_3(\gamma_1 + \gamma_2) + b(\gamma_0 - \gamma_3), \\ m_3 &= (\gamma_1 + \gamma_2)\, a + \gamma_2(\gamma_0 + \gamma_1 + \gamma_2 - \gamma_3) \end{aligned} \tag{A. 2}$$

where $\gamma_3 \leqq b \leqq \gamma_1 + \gamma_2 - \gamma_0$ and $m_1 < m_2 < m_3$ with $(m_1, m_2, m_3) = 1$. The curve C is the case $a = p$, $b = q$.

Case d). In this case we have an infinite family. The matrix A_2 of C has the form

$$\begin{pmatrix} X_2^{\gamma_2} & -X_0^{\gamma_0} X_3^{\gamma_3 - \gamma_{31}} & -X_1^t \\ -X_3^{\gamma_{31}} & X_1^{\gamma_1} & X_2^{t+c} X_0^{(n-4)\gamma_0 + \gamma_{0n-3}} \end{pmatrix} \tag{A. 3}$$

where $c = \gamma_{31} - \gamma_2 - (n-4)\,\gamma_0 - \gamma_{0n-3}$. Thus $\gamma_1, \gamma_2, \gamma_0, \gamma_3, \gamma_{31}, \gamma_{0n-3}, c$, and t may be all calculated. Then $C(m_1, m_2, m_3)$ is in the even liaison class of C if and only if

$$\begin{aligned} m_1 &= \gamma_3 a + \gamma_3 c + \gamma_2 \gamma_3 - \gamma_2 \gamma_{31}, \\ m_2 &= \gamma_3 a + \gamma_{31} \gamma_1, \\ m_3 &= (\gamma_1 + \gamma_2)\, a + \gamma_1(c + \gamma_2) \end{aligned} \tag{A. 4}$$

where $1 \leqq m_1 < m_2 < m_3$, $(m_1, m_2, m_3) = 1$ and a is arbitrary under the two conditions above. The curve C is the case $a = t$.

Case e). This case is the finite liaison class case. We know (see Theorem 4.1, Lemma 4.7) that A_2 has the form

$$\begin{pmatrix} X_3^{\gamma_3} & -X_0^{\gamma_0} X_2^{\gamma_2 - \gamma_{21}} & -X_1^{(n-4)\gamma_1 + \gamma_{1n-3}} \\ -X_2^{\gamma_{21}} & X_1^{\gamma_1} & X_3^s X_0^{(n-4)\gamma_1 + \gamma_{1n-3} + \gamma_{21} - s - \gamma_3} \end{pmatrix}. \tag{A. 5}$$

Hence from C we may calculate $\gamma_1, \gamma_2, \gamma_3, \gamma_{21}, \gamma_{1n-3}$, and s from the matrix A_2.

Then the even liaison class of C consists of all $C(m_1, m_2, m_3)$ such that

$$\begin{aligned} m_1 &= \gamma_2 b + \gamma_3(\gamma_2 - \gamma_{21}), \\ m_2 &= \gamma_1 b + ((n-3)\,\gamma_1 + \gamma_{1n-3})\,\gamma_3, \\ m_3 &= ((n-4)\,\gamma_1 + \gamma_{1n-3})\,\gamma_2 + \gamma_1 \gamma_{21}, \end{aligned} \tag{A. 6}$$

where $m_1 < m_2 < m_3$ and $(m_1, m_2, m_3) = 1$ and where b ranges between

$$0 \leqq b \leqq (n-4)\,(\gamma_1 - \gamma_0) + \gamma_{1n-3} + \gamma_{21} - \gamma_3 - \gamma_0.$$

The curve C is the case $b = s$.

Example. We close with two examples to illustrate how this appendix may be used. Consider $C = C(5, 8, 9)$. The set

$$S_2 = \{X_1^5 - X_2^2 X_3 X_0^2,\ X_1^3 X_3 - X_0 X_2^3,\ X_1^2 X_2 - X_0 X_3^2\}.$$

The dual curve with $X_3=1$ is parametrized by $1, 4, 9$ (with coordinates X_2, X_1, X_0) and hence

$$S_{n-1}=\{X_1X_3^3-X_2^4,\ X_3X_1^3-X_0X_2^3,\ X_1^2X_2-X_0X_3^2\}.$$

Hence $|S_2\cap S_{n-1}|=2$ and we are in case b). Writing the elements of S_2 as the 2×2 minors of A_2 we obtain

$$A_2=\begin{pmatrix} X_3 & -X_0X_2^2 & -X_1^2 \\ -X_2 & X_1^3 & X_0X_3 \end{pmatrix}.$$

Comparing with (A. 1) yields $\gamma_1=2$, $\gamma_2=1$, $\gamma_3=1$, $\gamma_0=1$. Hence the even liaison class is given by $C(ab+b-1, ab+b+2, 3a+3)$ where $1\leqq b\leqq 2$, or alternatively is the union of

$$\{C(a, a+3, 3a+3)\mid 1\leqq a,\ (a,3)=1\}\cup\{C(2a+1, 2a+4, 3a+3)\mid 2\leqq a,\ a\not\equiv 1 \bmod 3\}.$$

$C(5, 8, 9)$ is the case where $a=2$, $b=2$. (See example 4. 28 also.)

Example. For a different example consider $C(5, 19, 29)$. The set

$$S_2=\{-X_0^{14}X_2^5+X_1^{19},\ -X_0^2X_3+X_1^2X_2,\ -X_0^{12}X_2^6+X_1^{17}X_3\}$$

while

$$S_{n-1}=\{X_2^{12}-X_1^5X_3^7,\ -X_0^2X_2^{11}+X_1^7X_3^6,\ -X_0^2X_3+X_1^2X_3\}.$$

Hence $S_2\cap S_{n-1}=\{X_0^2X_3-X_1^2X_2\}$ and we are in case d). The matrix A_2 can be claculated:

$$A_2=\begin{pmatrix} X_2 & -X_0^2 & -X_1^{17} \\ -X_3 & X_1^2 & X_2^5X_0^{12} \end{pmatrix},$$

comparing with (A. 3) yields $\gamma_0=2$, $\gamma_1=2$, $\gamma_2=1$, $\gamma_{31}=\gamma_3=1$, $c=-12$. Hence the even linkage class is given by $\{C(a-12, a+2, 3a-22)\}$ whenever $1\leqq a-12<a+2<3a-22$ and these three numbers are relatively prime. Hence $a\geqq 13$ and $(a, 2)=1$ and $a\not\equiv 5 \bmod 7$. The curve $C(5, 19, 29)$ is the case where $a=17$.

Acknowledgement. We would like to express our appreciation to Professor W. Vogel for an invitation to the Martin-Luther Universität in Halle, G.D.R. in 1983 which made it possible to initiate our work on this paper.

References

[1] *H. Bresinsky*, Minimal free resolutions of monomial curves in $\mathbb{P}_k^3$, Linear Alg. Appl. **59** (1984), 121—129.
[2] *H. Bresinsky*, On the Cohen-Macaulay property for monomial curves in $\mathbb{P}_k^3$, Monatshefte für Math. **98** (1984), 21—28.
[3] *D. Buchsbaum* and *D. Eisenbud*, What makes a complex exact? J. Alg. **25** (1973), 259—268.
[4] *J. Herzog*, Note on complete intersections, in: E. Kunz, Einführung in die kommutative Algebra und algebraische Geometrie, Braunschweig-Wiesbaden 1979, 142—144.
[5] *C. Peskine* and *L Szpiro*, Liaison des varietés algébriques, Invent. Math. **26** (1973), 271—302.
[6] *P. Rao*, Liaison among curves in $\mathbb{P}_k^3$, Invent. Math. **50** (1979), 205—217.
[7] *P. Schwartau*, Liaison addition and monomial ideals, Thesis, Brandeis 1981.

Department of Mathematics, University of Maine, Orono, ME 04469, USA

Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA

Eingegangen 23. November 1984

On the definition and properties of p-superharmonic functions

By *Peter Lindqvist* at Espoo

1. Introduction

The solutions of the partial differential equation

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u)=0, \tag{1.1}$$

$1<p<\infty$, are called p-harmonic functions. They form a similar basis for a prototype of a non-linear potential theory as harmonic functions do in the corresponding classical theory. Especially, the celebrated method of O. Perron can be applied even in the non-linear situation $p\neq 2$. See [6] for this method and related results concerning more general elliptic equations than (1.1).

In this connexion the so called *p-superharmonic functions* play a rôle analogous to that of ordinary superharmonic functions in Potential Theory. Imitating the classical definition of F. Riesz, we say that a lower semi-continuous function $v: G\to(-\infty,\infty]$ is p-superharmonic in the domain $G\subset\mathbb{R}^n$, if v obeys the comparison principle with respect to p-harmonic functions (the case $v\equiv\infty$ being excluded); see Definition 2.2.

In literature p-superharmonic functions have been treated as weak *supersolutions* of (1.1). These are functions v belonging locally to Sobolev's space $W_p^1(G)$ and satisfying the inequality

$$\int|\nabla v|^{p-2}\nabla v\cdot\nabla\eta\,dm\geqq 0, \tag{1.2}$$

whenever $\eta\in C_0^\infty(G)$ is non-negative. The requirement that supersolutions belong to $\operatorname{loc}W_p^1(G)$ is necessary for (1.2) to make sense. However, from a potential theoretic point of view such an assumption is too restrictive and not quite adequate. For example, "the fundamental solutions"

$$v(x)=\int_{|x|}^{1} t^{-\frac{n-1}{p-1}}dt \tag{1.3}$$

do not satisfy this assumption in the ball $|x|=(x_1^2+\cdots+x_n^2)^{\frac{1}{2}}<1$. But according to our definition they are p-superharmonic in the unit ball, indeed.

The price that one has to pay for dropping all assumptions on the derivatives and, consequently, for replacing (1. 2) by the comparison principle is the seemingly much weaker requirement that p-superharmonic functions be lower semi-continuous by definition. It is remarkable that the delicate gap between the concepts "supersolution" and "p-superharmonic function" can be bridged over to some extent.

One of the main points in this kind of non-linear potential theory is that *bounded* p-superharmonic functions, when all comes to all, belong to $\operatorname{loc} W_p^1(G)$ and that (1. 2) is relevant! This question has been treated in [6] and we shall give only a brief review of the bounded case. — The purpose of our article is to study unbounded p-superharmonic functions.

At the focus of our attention is merely the general behaviour of the functions themselves and not the theory of p-polar set, i.e., not the structure itself of those sets where p-superharmonic functions attain the value $+\infty$. A central result of this genre is the theorem below. Its proof is given in Chapter 4.

1. 4. Theorem. *Suppose that v is p-superharmonic in the domain $G \subset \mathbb{R}^n$, $p > 2 - \frac{1}{n}$. Then the Sobolev derivative $\nabla v = \left(\frac{\partial v}{\partial x_1}, \ldots, \frac{\partial v}{\partial x_n} \right)$ exists and the local integrability result*

$$\int_D |\nabla v|^q \, dm < \infty \qquad (D \Subset G) \tag{1. 5}$$

holds, whenever $0 < q < \frac{n(p-1)}{n-1}$.

The functions (1. 3) show that the bound for the exponent q in (1. 5) is sharp. The restriction $p > 2 - \frac{1}{n}$ is not essential but in the cases $1 < p \leqq 2 - \frac{1}{n}$ the interpretation of ∇v is, in general, more involved (Remark 4. 12).

In Theorem 5. 4 we show that

$$v(x) = \lim_{y \to x} \operatorname{essinf} v(y)$$

and combining this result with (1. 5) we get a counterpart to Evans' celebrated result [3] on the ACL-properties (ACL = absolutely continuous along a.e. line) of superharmonic functions.

Finally, we mention that the results and methods can be readily extended to potential theories induced by more general elliptic partial differential equations, such as those considered in [8] and [6]. We have chosen the p-harmonic equation (1. 1) as a prototype just for instructive purposes. On the other hand, we think that some features in our exposition might be interesting even in the classical case $p = 2$, when (1. 1) reduces to Laplace's equation $\Delta u = 0$.

Notation. We use mainly standard notations. If $D \subset \mathbb{R}^n$ is an open set, $C(D)$ and $C(\bar{D})$ denote the class of continuous real valued functions in D and in the closure $\bar{D}$ of D, respectively. The symbol $C_0^\infty(D)$ denotes the class of infinitely many times differentiable functions with compact support in D. Sobolev's space $W_p^1(D)$ consists of those measurable functions $u: D \to \overline{\mathbb{R}}$ that together with their first generalized partial derivatives $\nabla u = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right)$ are p-summable in D. The corresponding local space is $\operatorname{loc} W_p^1(D)$. The closure of $C_0^\infty(D)$ in $W_p^1(D)$ is $W_{p,0}^1(D)$.

The open ball centered at x with radius r is denoted by $B(x, r)$.

2. Preliminaries

Suppose that G is a domain in $\mathbb{R}^n$. A function

$$h \in C(G) \cap \operatorname{loc} W_p^1(G)$$

is called *p-harmonic* in G, if

$$\int_G |\nabla h|^{p-2} \nabla h \cdot \nabla \eta \, dm = 0 \tag{2.1}$$

for every test-function $\eta \in C_0^\infty(G)$. A profound study of p-harmonic functions is given in [4] and [11]. See also [1] for some interesting connexions in Analysis.

If G is a bounded domain and $\varphi \in C(\bar{G}) \cap W_p^1(G)$, then there is a unique p-harmonic function h in G such that $h - \varphi \in W_{p,0}^1(G)$. If G is a *regular domain*, then $h|\partial G = \varphi|\partial G$ and $h \in C(\bar{G})$. For example, a ball is a regular domain. For our purpose it is sufficient to know that any domain G can be exhausted by regular domains. See [8] for this fact.

Harnack's principle holds in its classical version, i.e., if $h_1 \leqq h_2 \leqq \cdots$ all are p-harmonic in G, then $h = \lim h_k$ is either p-harmonic in G or identically $+\infty$ in G.

2.2. Definition. A function $v: G \to (-\infty, \infty]$ is called *p-superharmonic in G*, if

(i) v is lower semi-continuous in G,

(ii) $v \not\equiv \infty$ in G,

(iii) for each domain $D \Subset G$ the comparison principle holds: if $h \in C(\bar{D})$ is p-harmonic in D and $h|\partial D \leqq v|\partial D$, then $h \leqq v$ in D.

For $p = 2$ this is the classical definition of F. Riesz, c.f. [10], 1.2. We emphasize that not even the existence of the derivatives

$$\nabla v = \left(\frac{\partial v}{\partial x_1}, \ldots, \frac{\partial v}{\partial x_n}\right)$$

is required in Definition 2.2. However, for sufficiently regular p-superharmonic functions we have the following well-known characterization.

2. 3. Theorem. *Suppose that v belongs to $C(G) \cap \operatorname{loc} W_p^1(G)$. Then the following conditions are equivalent:*

(i) $\int_D |\nabla v|^p \, dm \leqq \int_D |\nabla(v+\eta)|^p \, dm$, *whenever* $D \Subset G$ *is a domain and* $\eta \in C_0^\infty(D)$ *is non-negative.*

(ii) $\int_G |\nabla v|^{p-2} \nabla v \cdot \nabla \eta \, dm \geqq 0$, *whenever* $\eta \in C_0^\infty(G)$ *is non-negative.*

(iii) v *is p-superharmonic in* G.

Proof. The equivalence of (i) and (ii) as well as the fact that (ii) implies (iii) can be found in any standard text such as [5] but we refer to [7], where also the sufficiency of (iii) is demonstrated. □

2. 4. Remarks. (1°) Conditions (i) and (ii) are equivalent for any $v \in \operatorname{loc} W_p^1(G)$.

(2°) See Corollary 3. 6 for a stronger version.

(3°) By an approximation it is easily seen that, if (i) holds, then (i) is true for every $\eta \geqq 0$ belonging to $W_{p,0}^1(D)$. A similar remark concerns (ii).

Theorem 2. 3 has some well-known consequences but for the sake of completeness we give concise proofs for them.

2. 5. Corollary. *Suppose that $v \in C(G) \cap \operatorname{loc} W_p^1(G)$ is p-superharmonic in G. If $D \Subset G$ is a domain, then*

$$\int_D |\nabla v|^p \, dm \leqq (p \operatorname*{osc}_{\operatorname{spt}\zeta} v)^p \int_G |\nabla \zeta|^p \, dm \tag{2. 6}$$

whenever $\zeta \in C_0^\infty(G)$, $0 \leqq \zeta \leqq 1$, $\zeta | D = 1$. *Here* $\operatorname{spt}\zeta$ *denotes the closure of the set* $\{x \mid \zeta(x) \neq 0\}$.

Proof. By Theorem 2. 3 (and Remark 2. 4 (3°))

$$\int_{\operatorname{spt}\zeta} |\nabla v|^p \, dm \leqq \int_{\operatorname{spt}\zeta} |\nabla(v + \zeta^p(M-v))|^p \, dm \tag{2. 7}$$

where $M = \sup_{\operatorname{spt}\zeta} v$. Now $\nabla(v + \zeta^p(M-v)) = (1-\zeta^p) \nabla v + p\zeta^{p-1}(M-v) \nabla \zeta$ and by convexity and homogenity

$$|\nabla(v + \zeta^p(M-v))|^p \leqq (1-\zeta^p) |\nabla v|^p + p^p |M-v|^p |\nabla \zeta|^p.$$

Integrating this inequality over $\operatorname{spt}\zeta$ and using (2. 7), we arrive at (2. 6) after some simplification. □

2. 8. Corollary. *Suppose that $v > 0$, $v \in C(G) \cap \operatorname{loc} W_p^1(G)$, and that v is p-superharmonic in G. Then*

$$\int_D |\nabla \log v|^p \, dm \leqq \left(\frac{p}{p-1}\right)^p \int_G |\nabla \zeta|^p \, dm \tag{2. 9}$$

whenever $D \Subset G$ and $\zeta \in C_0^\infty(G)$, $0 \leqq \zeta \leqq 1$, $\zeta | D = 1$.

Proof. Fix ζ. We may assume that $v \geqq (p-1)^{\frac{1}{p}}$ in the set where $\zeta \neq 0$, since (2.9) is invariant under the multiplication λv, $\lambda > 0$ being a constant. By Theorem 2.3 (and Remark 2.4 (3°))

$$\int_{\operatorname{spt}\zeta} |\nabla v|^p \, dm \leqq \int_{\operatorname{spt}\zeta} |\nabla(v + \zeta^p v^{-p+1})|^p \, dm. \tag{2.10}$$

Now $\nabla(v + \zeta^p v^{-p+1}) = \left[1 - (p-1)\left(\frac{\zeta}{v}\right)^p\right] \nabla v + p\left(\frac{\zeta}{v}\right)^{p-1} \nabla\zeta$ and by convexity

$$|\nabla(v + \zeta^p v^{-p+1})|^p \leqq \left(1 - \frac{(p-1)\,\zeta^p}{v^p}\right) |\nabla v|^p + \frac{p^p}{(p-1)^{p-1}} |\nabla\zeta|^p.$$

Integrating this inequality over $\operatorname{spt}\zeta$ and using (2.10) we arrive at the bound

$$(p-1) \int_{\operatorname{spt}\zeta} \zeta^p |\nabla \log v|^p \, dm \leqq \frac{p^p}{(p-1)^{p-1}} \int_{\operatorname{spt}\zeta} |\nabla\zeta|^p \, dm.$$

This yields (2.9). □

3. Approximation

The key to the study of functions that are p-superharmonic in the sense of Definition 2.2 is provided by a variational *obstacle problem.*

Fix any function $\varphi \in C^\infty(\mathbb{R}^n)$ and choose a bounded and regular domain $D \subset \mathbb{R}^n$. The problem of minimizing the variational integral $\int |\nabla u|^p \, dm$ in the class

$$\mathcal{F}_\varphi(D) = \{u \in C(\bar{D}) \cap W_p^1(D) \mid u \geqq \varphi \quad \text{in} \quad D, \; u|\partial D = \varphi|\partial D\}$$

leads to p-superharmonic functions. (The regularity of D is not essential but a dropping of this assumption would force us to consider the boundary values in the $W_{p,0}^1$-sense, an adjustment later causing minor technical complications.)

3.1. Lemma. *There is a unique function $v_\varphi \in \mathcal{F}_\varphi(D)$ such that*

$$\int_D |\nabla v_\varphi|^p \, dm \leqq \int_D |\nabla u|^p \, dm \tag{3.2}$$

whenever $u \in \mathcal{F}_\varphi(D)$. Moreover, v_φ is p-superharmonic in D and p-harmonic in (each component of) the open set $\{x \in G \mid v_\varphi(x) > \varphi(x)\}$.

Proof. The existence of a unique minimizing $v_\varphi \in \mathcal{F}_\varphi(D)$ is nowadays well-known[1]) in the Calculus of Variations, and according to Theorem 2.3 v_φ is p-superharmonic in D.

It remains for us to show that v_φ is p-harmonic in the set

$$A = \{x \in D \mid v_\varphi(x) > \varphi(x)\},$$

where the obstacle does not hinder. This set is open. If A is non-empty, fix any test-function $\eta \in C_0^\infty(A)$ such that $|\eta| \leqq 1$. Then $\kappa = \sup_{\eta \neq 0} (v_\varphi - \varphi)$ is positive and $v_\varphi + \varepsilon\eta \in \mathcal{F}_\varphi(D)$, whenever $|\varepsilon| \leqq \kappa$. For $|\varepsilon| \leqq \kappa$ we have

$$\int_D |\nabla v_\varphi|^p \, dm \leqq \int_D |\nabla(v_\varphi + \varepsilon\eta)|^p \, dm.$$

---

[1]) *Note added in proof.* See the methods in: J. Michel and P. Ziemer, Interior regularity for solutions to obstacle problems (to appear).

This implies, by a device originally due to Lagrange, that

$$\int_D |\nabla v_\varphi|^{p-2} \nabla v_\varphi \cdot \nabla \eta \, dm = 0. \tag{3.3}$$

The restriction $|\eta| \leqq 1$ is easily removed from (3.3) and so

$$\int_A |\nabla v_\varphi|^{p-2} \nabla v_\varphi \cdot \nabla \eta \, dm = 0$$

whenever $\eta \in C_0^\infty(A)$. This means that v_φ is p-harmonic in A. □

Via Lemma 3.1 a constructive approach to p-superharmonic functions is possible. Suppose that v is p-superharmonic in G and fix a regular domain $D \Subset G$. By the lower semi-continuity of v there are functions $\varphi_1 \leqq \varphi_2 \leqq \cdots$ in $C^\infty(\mathbb{R}^n)$ such that

$$v(x) = \lim_{k \to \infty} \varphi_k(x),$$

when $x \in \bar{D}$ [10], 1.3.

Let $v_k \in \mathcal{F}_{\varphi_k}(D)$ denote the p-superharmonic function v_{φ_k} in Lemma 3.1. Then $\varphi_k \leqq v_k$ and we claim that

$$v_1 \leqq v_2 \leqq \cdots, \quad v = \lim v_k$$

in D. To establish that $v = \lim v_k$ it is sufficient to show that $v_k \leqq v$ for $k = 1, 2, \ldots$. By Lemma 3.1, v_k is p-harmonic in the open set $A_k = \{v_k > \varphi_k\}$. Since $v_k | \partial A_k = \varphi_k | \partial A_k \leqq v | \partial A_k$, we have $v_k \leqq v$ in A_k by the comparison principle. This shows that $v_k \leqq v$ in D, the statement being selfevident in $D \setminus A_k$.

A similar argument shows that $v_k \leqq v_{k+1}$ in D.

3.4. Theorem. *Suppose that v is p-superharmonic in G. Given a domain $D \Subset G$, there are such p-superharmonic functions $v_k \in C(\bar{D}) \cap W_p^1(D)$ that $v_1 \leqq v_2 \leqq \cdots$ and $\lim v_k = v$ in D. If, in addition, v is locally bounded from above in G, then $v \in \operatorname{loc} W_p^1(G)$ and the approximants v_k can be chosen so that*

$$\lim_{k \to \infty} \int_D |\nabla (v - v_k)|^p \, dm = 0.$$

Proof. Fix $D \Subset G$ and choose a regular domain D_1, $\bar{D} \subset D_1 \Subset G$. By the previous construction there are p-superharmonic functions v_k in D_1 such that $v_k \to v$ pointwise in D_1 and $v_k \in C(D_1) \cap W_p^1(D_1)$. This proves the first part of the theorem.

If v is locally bounded from above in G, then

$$C = \sup_{D_1} v - \inf_{D_1} v_1 < \infty$$

and so Corollary 2.5 gives the bound

$$\int_B |\nabla v_k|^p \, dm \leqq p^p C^p \int_{D_1} |\nabla \zeta|^p \, dm = M_B \qquad (k = 1, 2, \ldots)$$

for any ball $B \Subset D_1$. By a standard argument $v \in W_p^1(B)$ and $\nabla v_k \to \nabla v$ weakly in $L^p(B)$. A finite number of such balls covers $\bar{D}$ and hence $v \in W_p^1(D)$. Since $D \Subset G$ was arbitrary, $v \in \operatorname{loc} W_p^1(G)$.

To establish the strong convergence of the derivatives it is enough to show that

$$\lim_{k\to\infty}\int_{B_r}|\nabla v-\nabla v_k|^p\,dm=0,$$

whenever B_r is such a ball with radius r that the closure of the concentric ball B_{2r} with radius $2r$ is in D_1. To this end, fix $\eta\in C_0^\infty(B_{2r})$, $0\leqq\eta\leqq 1$, $\eta\,|\,B_r=1$.

By Hölder's inequality the last integral in the expression

$$\begin{aligned}I_k&=\int[|\nabla v|^{p-2}\nabla v-|\nabla v_k|^{p-2}\nabla v_k]\cdot\nabla[\eta(v-v_k)]\,dm\\&=\int\eta[|\nabla v|^{p-2}\nabla v-|\nabla v_k|^{p-2}\nabla v_k]\cdot[\nabla v-\nabla v_k]\,dm\\&+\int(v-v_k)[|\nabla v|^{p-2}\nabla v-|\nabla v_k|^{p-2}\nabla v_k]\cdot\nabla\eta\,dm\end{aligned}$$

is bounded in absolute value by the quantity

$$\left\{\int_{B_{2r}}(v-v_k)^p\,dm\right\}^{\frac{1}{p}}\left\{\left(\int_{B_{2r}}|\nabla v|^p\,dm\right)^{1-\frac{1}{p}}+\left(\int_{B_{2r}}|\nabla v_k|^p\,dm\right)^{1-\frac{1}{p}}\right\}\cdot\max|\nabla\eta|.$$

Using the obvious fact that $\int_{B_{2r}}(v-v_k)^p\,dm\to 0$ as $k\to\infty$, we arrive at

$$(3.5)\qquad \overline{\lim_{k\to\infty}}\int\eta[|\nabla v|^{p-2}\nabla v-|\nabla v_k|^{p-2}\nabla v_k]\cdot[\nabla v-\nabla v_k]\,dm\leqq\overline{\lim_{k\to\infty}}\,I_k.$$

On the other hand Theorem 2.3 yields

$$I_k\leqq\int|\nabla v|^{p-2}\nabla v\cdot\nabla[\eta(v-v_k)]\,dm$$

and $\overline{\lim}\,I_k\leqq 0$, since $\nabla\eta v_k\to\nabla\eta v$ weakly in $L^p(B_{2r})$. According to (3.5) we have

$$\overline{\lim_{k\to\infty}}\int\eta[|\nabla v|^{p-2}\nabla v-|\nabla v_k|^{p-2}\nabla v_k]\cdot[\nabla v-\nabla v_k]\,dm\leqq 0.$$

Because $[|y|^{p-2}y-|x|^{p-2}x]\cdot(y-x)\geqq 2^{1-p}|y-x|^p$ for $x,y\in\mathbb{R}^n$ we arrive at

$$\lim_{k\to\infty}\int_{B_{2r}}\eta|\nabla v-\nabla v_k|^p\,dm=0$$

from which the desired strong convergence follows. □

3.6. Corollary. *Suppose that v is p-superharmonic and locally bounded from above in G. Then $v\in\operatorname{loc}W_p^1(G)$ and*

$$\int|\nabla v|^{p-2}\nabla v\cdot\nabla\eta\,dm\geqq 0,$$

whenever $\eta\in C_0^\infty(G)$ is non-negative.

Proof. By Theorem 3.4 and Theorem 2.3

$$\int|\nabla v|^{p-2}\nabla v\cdot\nabla\eta\,dm=\lim_{k\to\infty}\int|\nabla v_k|^{p-2}\nabla v_k\cdot\nabla\eta\,dm\geqq 0$$

whenever $\eta\in C_0^\infty(G)$, $\eta\geqq 0$. □

3.7. Remark. By means of the described approximation process estimates such as (2.6) and (2.9) are readily extended to locally bounded p-superharmonic functions, whether these are continuous or not.

4. Integrability

If the function v is p-superharmonic in G, so are the functions

$$v_k = \min\{v, k\} \qquad (k = 1, 2, \dots).$$

Because v_k is locally bounded in G, ∇v_k exists and $v_k \in \operatorname{loc} W_p^1(G)$ by Theorem 3.4.[2]) We aim at studying those features that are somehow preserved under the limit process

$$v = \lim_{k \to \infty} v_k.$$

(Only the cases $1 < p \leqq n$ are interesting from this point of view, because any p-superharmonic function is continuous, if $p > n$.)

Our first step is to show that the set where $v < +\infty$ is dense in G.

4.1. Proposition. *If v is p-superharmonic in G, then the set where $v = +\infty$ does not contain any open ball.*

Proof. Suppose to begin with that $v > 0$ in G. Assume that $v \equiv +\infty$ in some ball $B_r = B(x_0, r) \Subset G$. Choose $R > r$ so that $B_R = B(x_0, R) \Subset G$. Consider the ring domain $B_R \setminus \bar{B}_r$. There is a (unique) p-harmonic function h_k in $B_R \setminus \bar{B}_r$ with boundary values $h_k | \partial B_R = 0$, $h_k | \partial B_r = k$, where $k = 1, 2, \dots$. To be more explicit,

$$h_k(x) = k \frac{\int_{|x - x_0|}^{R} t^{-\frac{n-1}{p-1}} dt}{\int_{r}^{R} t^{-\frac{n-1}{p-1}} dt} \qquad (k = 1, 2, \dots)$$

and $h_k = k h_1$.

Now $v \geqq h_k$, $k = 1, 2, \dots$, on the boundary of $B_R \setminus \bar{B}_r$ and hence in that domain itself. This means that $v \geqq \lim h_k = +\infty$ everywhere in $B_R \setminus \bar{B}_r$. Thus $v \equiv +\infty$ in B_R.

To get rid of the restriction $v > 0$, we consider the function $v - \inf_{B_R} v$ instead of v in the previous construction and again we obtain that $v | B_R \equiv \infty$, if $v | B_r \equiv \infty$.

Repeating this process with a suitable chain of balls we finally arrive at the contradiction $v \equiv +\infty$ in G, a case that was excluded by Definition 2.2. □

The functions (1.3) show that the degree of integrability is sharp in the theorem below.

4.2. Theorem. *If v is p-superharmonic in $G \subset \mathbb{R}^n$, $1 < p \leqq n$, then*

$$\int_D |v|^q \, dm < +\infty \tag{4.3}$$

whenever $D \Subset G$ and $0 \leqq q < \dfrac{n(p-1)}{n-p}$.

---

[2]) v_k has a different meaning here than in Chapter 3.

Proof. Because of the local nature of (4. 3) we may assume that $v>0$. Then also $v_k>0$, where $v_k=\min\{v, k\}$, $k=1, 2, \ldots$. By Corollary 3. 6 each v_k satisfies the inequality

$$\int |\nabla v_k|^{p-2}\nabla v_k\cdot\nabla\eta\, dm\geqq 0, \quad (k=1, 2, \ldots) \tag{4. 4}$$

whenever $\eta\in C_0^\infty(G)$, $\eta\geqq 0$. According to a profound result of Trudinger, c.f. [12], for differential inequalities such as (4. 4) a bound of the type

$$\Big(\int_{B_r} v_k^q dm\Big)^{\frac{1}{q}}\leqq C(n, p, q, r)\cdot \operatorname*{essinf}_{B_r} v_k \quad (k=1, 2, \ldots) \tag{4. 5}$$

holds, where $C(n, p, q, r)<\infty$ for $q<\dfrac{n(p-1)}{n-p}$ is independent of v_k. Here B_r is any ball with radius r such that the concentric ball with radius $4r$ is in G.

(Trudinger's proof is based on the lemma of John and Nirenberg. The proof can be simplified in the favourable case $p=n$ and the case $p=2$ is illuminated by [5], Theorem 8. 18.)

By Proposition 4. 1 and Theorem 5. 4[3])

$$\operatorname*{essinf}_{B_r} v_k\leqq \operatorname*{essinf}_{B_r} v<\infty$$

and so Fatou's lemma together with (4. 5) gives

$$\int_{B_r} v^q dm<\infty.$$

Now a suitable covering of the compact set $\bar{D}$ with balls and the above bound give (4. 3). □

By the observation in Remark 3. 7

$$\int_D |\nabla\log v_k|^p\, dm\leqq\left(\frac{p}{p-1}\right)^p\int_G |\nabla\zeta|^p\, dm, \quad (k=1, 2, \ldots) \tag{4. 6}$$

when $D\Subset G$ provided that $v>0$ in G. By the aids of (4. 6), we immediately get the following result.

4. 7. Theorem. *Suppose that v is p-superharmonic and positive in G. Then Sobolev's derivative $\nabla\log v$ exists and* (2. 9) *holds.*

The obtained local bound for $\int|\nabla\log v|^p\, dm$, $v>0$, implies by well-known properties of Sobolev's space that every p-superharmonic function with $p>n$ is continuous and hence locally bounded. (No redefinition of v in a set of Lebesgue measure zero is performed to reach this continuity: the original function itself is continuous by Theorem 5. 4 below.) Thus only the cases $1<p\leqq n$ are of actual interest. — By the way, we mention that $\log v$ is p-superharmonic in G.

For the proof of Theorem 1. 4 we need an estimate of auxiliary nature.

4. 8. Lemma. *Suppose that $v\geqq 1$ and p-superharmonic in G. Then*

$$\int_D v_k^{-1-\alpha}|\nabla v_k|^p\, dm\leqq\frac{p^p}{\alpha^p}\int_G v_k^{p-1}|\nabla\zeta|^p\, dm \tag{4. 9}$$

whenever $D\Subset G$ is a domain, $\alpha>0$, $\zeta\in C_0^\infty(G)$, $0\leqq\zeta\leqq 1$, and $\zeta|D=1$.

---

[3]) The proof for Theorem 5. 4 does not depend on Chapter 4.

Proof. Suppose to begin with that $0<\alpha\leqq 1$. According to Corollary 3.6 and Remark 2.4 (1°)

$$\int_{\operatorname{spt}\zeta} |\nabla v_k|^p \, dm \leqq \int_{\operatorname{spt}\zeta} |\nabla (v_k + \zeta^p v_k^{-\alpha})|^p \, dm.$$

Using the same device as in the proof of Corollary 2.8 we get

$$\alpha \int \zeta^p v_k^{-1-\alpha} |\nabla v_k|^p \, dm \leqq \alpha \left(\frac{p}{\alpha}\right)^p \int v_k^{p-1-\alpha} |\nabla \zeta|^p \, dm \tag{4.10}$$

for $k=1, 2, \ldots$. The restriction $\alpha \leqq 1$ was forced by the requirement that $\alpha \zeta^p v_k^{-1-\alpha} \leqq 1$ in connexion with the use of convexity in the proof for (4.10). However, (4.10) is invariant when ζ is replaced by $\varepsilon\zeta$, ε being a positive constant. Thus (4.10) holds for every $\alpha>0$.

Now $v_k^{p-1-\alpha} \leqq v_k^{p-1}$ and the desired estimate (4.9) follows. □

We are now in the position of proving Theorem 1.4. The proof is basically a combination of (4.3) and (4.9).

Proof for Theorem 1.4. Suppose first that $v \geqq 1$ in G. Fix $D \Subset G$ and $q < \dfrac{n(p-1)}{n-1}$. By Hölder's inequality

$$\int_D |\nabla v_k|^q \, dm = \int_D v_k^{\frac{(1+\alpha)q}{p}} \left| \frac{\nabla v_k}{v_k^{\frac{1+\alpha}{p}}} \right|^q dm$$

$$\leqq \left\{ \int_D v_k^{\frac{(1+\alpha)q}{p-q}} dm \right\}^{1-\frac{q}{p}} \left\{ \int_D v_k^{-1-\alpha} |\nabla v_k|^p \, dm \right\}^{\frac{q}{p}}$$

for any $\alpha>0$. Note that $\dfrac{q}{p-q} < \dfrac{n(p-1)}{n-p}$. Hence we can fix $\alpha>0$ so that $\dfrac{(1+\alpha)q}{p-q} < \dfrac{n(p-1)}{n-p}$. For example

$$1+\alpha = \frac{\dfrac{q}{p-q} + \dfrac{n(p-1)}{n-p}}{\dfrac{2q}{p-q}}$$

will do for $1<p<n$ and any $\alpha>0$ is suitable for $p=n$. With this choice of α (4.9) yields

$$\int_D |\nabla v_k|^q \, dm \leqq \left\{ \int_D v_k^{\frac{(1+\alpha)q}{p-q}} dm \right\}^{1-\frac{p}{q}} \left(\frac{p}{\alpha}\right)^q \left\{ \int_G v_k^{p-1} |\nabla \zeta|^p \, dm \right\}^{\frac{q}{p}}$$

$$\leqq \left(\frac{p}{\alpha}\right)^q \sup |\nabla \zeta|^q \left\{ \int_D v^{\frac{(1+\alpha)q}{p-q}} dm \right\}^{1-\frac{p}{q}} \left\{ \int_{\operatorname{spt}\zeta} v^{p-1} dm \right\}^{\frac{q}{p}}$$

and so the sequence $\int_D |\nabla v_k|^q \, dm$, $k=1, 2, \ldots$, is uniformly bounded in virtue of (4.3).

Now a standard reasoning shows that Sobolev's derivative ∇v exists in D and

$$\int_D |\nabla v|^q \, dm \leqq \varliminf_{k \to \infty} \int_D |\nabla v_k|^q \, dm. \tag{4. 11}$$

Since $D \Subset G$ was arbitrary, we have $v \in \operatorname{loc} W_q^1(G)$.

The restriction that $v \geqq 1$ is again locally removed by a translation with a constant. This concludes our proof. □

4. 12. Remark. For $1 < p \leqq 2 - \frac{1}{n}$ the bound $\frac{n(p-1)}{n-1} \leqq 1$ for the integrability exponent q in $\int |\nabla v_k|^q \, dm$. Thus the derivative ∇v does not, in general, exist in the sense of Sobolev, i.e. the formula

$$\int \eta \nabla v \, dm = -\int v \nabla \eta \, dm,$$

$\eta \in C_0^\infty(G)$, cannot be used. However, ∇v can be given a suitable interpretation via the functions $\min\{v, k\}$, $k = 1, 2, \ldots$. Another approach is provided by the interpretation $\nabla v = v \nabla \log v$, $v > 0$, $\nabla \log v$ being the Sobolev derivative in Theorem 4. 7.

5. ACL-properties

If v is p-superharmonic in G, then for every $x \in G$

$$v(x) \leqq \varliminf_{y \to x} v(y) \leqq \operatorname{ess} \varliminf_{y \to x} v(y) \tag{5. 1}$$

by lower semicontinuity. Here essential limes inferior means that any set of Lebesgue n-measure zero can be neglected, when limes inferior is calculated. The precise definition of this concept is given in [2], II. 5. The reversed inequality is less obvious and its proof is divided in several steps.

5. 2. Lemma. *Suppose that v is p-superharmonic in G. If $v(x) \leqq 0$ for each $x \in G$ and if $v(x) = 0$ for a.e. $x \in G$, then $v(x) = 0$ for each $x \in G$.*

Proof. Choose $x \in G$. We claim that $v(x) = 0$. To this end, fix a domain D and a ball B so that $x \in B \Subset D \Subset G$. Let $v_1 \leqq v_2 \leqq \cdots$ be the approximants in Theorem 3. 4 for v in D, i.e. $\lim v_k(x) = v(x)$ for each $x \in D$ and $v_k \in C(D) \cap W_p^1(D)$.

In B there is a unique p-harmonic function $h_k \in C(\bar{B}) \cap W_p^1(B)$ such that $h_k | \partial B = v_k | \partial B$. It is easily verified that the so called Poisson modifications

$$V_k = \begin{cases} h_k & \text{in } B, \\ v_k & \text{in } D \setminus B, \end{cases}$$

are p-superharmonic in D. Moreover, $V_k \leqq v_k$ in D by the comparison principle.

Since $V_1 \leqq V_2 \leqq \cdots$ in D, $\lim V_k = V$ exists in D. Obviously, V is p-superharmonic in D (Definition 2. 2) and by Harnack's principle V is p-harmonic in B. By the construction $V = \lim V_k \leqq \lim v_k = v$ everywhere in D. Thus it is sufficient to show that $V(x) = 0$.

If $D_1 \Subset D$ is a domain such that $B \subset D_1$, then the construction and the proof for Corollary 2.5 yields

$$\underline{\lim} \int_{D_1} |\nabla V_k|^p \, dm \leqq \underline{\lim} \int_{D_1} |\nabla v_k|^p \, dm$$
$$\leqq p^p \underline{\lim} \int_D |v_k|^p \, |\nabla \zeta|^p \, dm = p^p \int_D |v|^p \, |\nabla \zeta|^p \, dm$$

where the last step follows from Lebesgue's convergence theorem. Clearly, ∇V exists and

$$\int_{D_1} |\nabla V|^p \, dm \leqq \underline{\lim} \int_{D_1} |\nabla V_k|^p \, dm = 0,$$

since

$$\int_{D_1} |\nabla V_k|^p \, dm \geqq \int_{D_1} |\nabla V|^p \, dm + p \int_{D_1} |\nabla V|^{p-2} \nabla V \cdot \nabla (V_k - V) \, dm$$

by convexity. Here the fact that $v = 0$ a.e. was used.

Thus $\nabla V = 0$ a.e. in D_1 and so V is constant a.e. in D_1. Note that $V = v$ everywhere in $D \setminus B$ by the construction. In particular, $V = 0$ a.e. in $D_1 \setminus B$. This means that $V = 0$ a.e. in D_1. Especially, $V = 0$ a.e. in B. But V, being p-harmonic in B, is continuous in B. Hence $V = 0$ everywhere in B. This contains the desired result $V(x) = 0$; remember that $0 = V(x) \leqq v(x) \leqq 0$. □

5.3. Lemma. *If v is p-superharmonic in G and if $v(x) > \lambda$ for a.e. $x \in G$, then $v(x) \geqq \lambda$ for every $x \in G$.*

Proof. (If $\lambda = -\infty$, there is nothing to prove.) Applying Lemma 5.2 to the p-superharmonic function defined by

$$\min \{v(x), \lambda\} - \lambda,$$

when $x \in G$, we obtain the desired result in the case $\lambda > -\infty$. □

The classical counterpart to the important phenomenon described in the next theorem seems to be credited to Brelot, c.f. [2], II § 5.

5.4. Theorem. *If v is p-superharmonic in G, then*

$$v(x) = \operatorname{ess} \underline{\lim}_{y \to x} v(y)$$

for each $x \in G$.

Proof. Fix any $x \in G$. According to (5.1) we must show that

$$\lambda = \operatorname{ess} \underline{\lim}_{y \to x} v(y) \leqq v(x).$$

Given any $\varepsilon > 0$, there is a radius $r_\varepsilon > 0$ such that $v(y) > \lambda - \varepsilon$ for *a.e.* $y \in B(x, r_\varepsilon)$, where $B(x, r_\varepsilon) \subset G$. Then $v(y) \geqq \lambda - \varepsilon$ for every $y \in B(x, r_\varepsilon)$ by Lemma 5.3. In particular $v(x) \geqq \lambda - \varepsilon$. Because $\varepsilon > 0$ was arbitrary, we have established that $\lambda \leqq v(x)$. □

In order to appreciate the following ACL-investigation, one should for example recall that given p, $1<p\leqq n$, there is a p-superharmonic function v in $\mathbb{R}^n$ such that $v(x)=+\infty$ if and only if the coordinates of x all are rational. (The construction of such a "monster" is immaterial in this connexion.) Moreover, things can be arranged so that $\int|\nabla v|^p\,dm<+\infty$ whenever the set of integration is bounded!

Recall the following concept. A function defined in a closed cube Q is of class ACL in Q if it is absolutely continuous along almost every line parallel to the sides. A function $v: G\to[-\infty,\infty]$ is of class ACL in G, if the restriction $v|Q$ is of class ACL in Q, whenever $Q\subset G$ is a closed cube. See [9], Chapter 3.

5. 5. Theorem (Evans). *Suppose that v is p-superharmonic in G, $p>2-\frac{1}{n}$. Then v is of class ACL in G.*

Proof. If $p>2-\frac{1}{n}$, then $\int_D|\nabla v|^q\,dm<\infty$, $D\Subset G$, for some fixed exponent $q>1$ (Theorem 1. 4). According to [9], Lemma 3. 1. 1 there is a function v^* that is of class ACL in G and $v^*=v$ a.e. in G. Using Theorem 5. 4 we get

$$v(x)=\operatorname{ess}\varliminf_{y\to x}v^*(y)$$

for every $x\in G$. Thus v coincides with v^* at least on every ACL-line of v^*. This means that v itself is of class ACL in G. □

References

[1] *B. Bojarski* and *T. Iwaniec*, p-Harmonic Equation and Quasiregular Mappings, Preprint **617**, Bonn 1983.

[2] *M. Brelot*, Éléments de la Théorie Classique du Potential (2e édition), Paris 1961.

[3] *G. Evans*, Complements of Potential Theory. II, Am. J. Math. **55** (1933), 29—49.

[4] *L. Evans*, A new proof of local $C^{1,\alpha}$ regularity for solutions of certain degenerate elliptic P.D.E., J. Diff. Equ. **45** (1982), 356—373.

[5] *D. Gilbarg* and *N. Trudinger*, Elliptic Partial Differential Equations of Second Order, Berlin-Heidelberg-New York 1977.

[6] *S. Granlund, P. Lindqvist* and *O. Martio*, Note on the PWB-method in the non-linear case (to appear in Pacific J. Math.).

[7] *P. Lindqvist*, On the comparison principle in the calculus of variations, Arkiv för mat. **21** (1983), 185—190.

[8] *V. Maz'ja*, On the continuity at a boundary point of solutions of quasilinear elliptic equations (in Russian), Vestnik Leningradskogo Universiteta **13** (1970), 42—55.

[9] *Ch. Morrey*, Multiple Integrals in the Calculus of Variations, Berlin-Heidelberg-New York 1966.

[10] *T. Radó*, Subharmonic Functions, New York 1949.

[11] *P. Tolksdorf*, On the Dirichlet problem for quasilinear equations in domains with conical boundary points (to appear in Communications in Partial Differential Equations).

[12] *N. Trudinger*, On Harnack type inequalities and their application to quasilinear elliptic equations, Comm. Pure Applied Math. **20** (1967), 721—747.

Helsinki University of Technology, Institute of Mathematics, SF-02150 Espoo, Finland

Eingegangen 10. Dezember 1984

On the reduction of the Poincaré rank of singular systems of ordinary differential equations

By *R. Schäfke and H. Volkmer* at Essen

1. Introduction.

We consider a system of linear differential equations in $\mathbb{C}^n$

$$z^{s+1}y' = A(z)y, \tag{1.1}$$

where $A(z)$ denotes an n by n matrix function that can be expanded in a power series at $z=0$ and s is a nonnegative integer. If $A(0)\neq 0$ then s is called the *Poincaré rank* of (1.1). If the Poincaré rank is positive then the behavior of the solutions near $z=0$ can be very complicated [1].

A transformation $y=T(z)\tilde{y}$ with $\det T(z)\not\equiv 0$ for $z\neq 0$ takes equations (1.1) into an equivalent system of differential equations

$$z^{s+1}\tilde{y}' = \tilde{A}(z)\tilde{y}, \text{ where } \tilde{A}(z)=T(z)^{-1}[A(z)T(z)-z^{s+1}T'(z)]. \tag{1.2}$$

If $T(z)$ is analytic at $z=0$ and $\det T(0)\neq 0$ the transformation is called *analytic*, if $T(z)$ (and then also $T(z)^{-1}$) has at most a pole at $z=0$ it is called *meromorphic*. Analytic transformations do not change the Poincaré rank, but simple examples show that meromorphic transformations can do so.

The growth order of the solutions of (1.1) as z tends to 0 is not greater than the Poincaré rank (see e.g. [7], p. 109) and it is invariant with respect to meromorphic transformations. This suggests the following questions:

1. Are there necessary and sufficient conditions in terms of $A(z)$ for the reducibility of the Poincaré rank of (1.1) by meromorphic transformations?

2. How can the minimal Poincaré rank of systems meromorphically equivalent to (1.1) be characterized?

3. How can one find a meromorphic transformation that reduces the Poincaré rank to the minimal one?

J. Moser [6] partially answered these questions. He proved: *The rank of* $A(0)$ *can be reduced by a meromorphic transformation if and only if*

$$z^{\operatorname{rank} A(0)} \det\left(\lambda I + z^{-1} A(z)\right)\big|_{z=0} = 0 \quad \textit{identically in } \lambda.$$

J. Moser also described how such a transformation can be found (V. Dietrich [2] improved his method and made it constructive throughout). By repeated application of the theorem one can check whether or not $\tilde{A}(0) = 0$ can be attained and thus the Poincaré rank can be reduced by one. Recursively one eventually finds the minimal Poincaré rank and a corresponding transformation. The disadvantage of this procedure is that only in the first step the criterion can be applied to $A(z)$ itself and that many steps may be necessary.

D. Lutz and R. Schäfke [5] answered the first and second question in the following way.

Let

$$A(z) = \sum_{k=0}^{\infty} A_k z^k$$

and for integers $m \geqq s$, $0 \leqq r \leqq s-1$ form the block matrices

$$\mathbf{A}_m^{(r)}(\lambda) := \begin{bmatrix}
A_0 & 0 & & & & \\
A_1 & A_0 & & & & \\
\vdots & A_1 & \ddots & & 0 & \\
A_{s-r-1} & \vdots & \ddots & \ddots & & \\
A_{s-r} - \lambda \cdot I & A_{s-r-1} & \ddots & \ddots & & \\
A_{s-r+} & A_{s-r} - \lambda I & \ddots & \ddots & \ddots & \\
\vdots & A_{s-r+1} & \ddots & \ddots & \ddots & \ddots \\
A_{s-1} & \vdots & \ddots & \ddots & \ddots & \ddots \\
A_s & A_{s-1} & \ddots & \ddots & \ddots & \ddots \\
A_{s+1} & A_s - I & \ddots & \ddots & \ddots & \ddots \\
\vdots & A_{s+1} & \ddots & \ddots & \ddots & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
A_{m-1} & A_{m-2} \ldots A_{s+1} A_s - (m-s-1) \cdot I & A_{s-1} \ldots & A_{s-r} - \lambda \cdot I \ldots & A_0 &
\end{bmatrix}.$$

All blocks in $\mathbf{A}_m^{(r)}$ on the same subdiagonal are equal except for those on the s^{th} subdiagonal which are equal to $A_s, A_s - I, \ldots, A_s - (m-s-1) \cdot I$. All blocks are constant except for the $(s-r)^{\text{th}}$ subdiagonal, whose entries equal $A_{s-r} - \lambda \cdot I$. Then they proved:

The system (1. 1) *is meromorphically equivalent to a system of Poincaré rank less or equal to* r *if and only if*

$$\operatorname{rank}_{C(\lambda)} \mathbf{A}_m^{(r)}(\lambda) = n(m-s+r) \quad \textit{for some} \quad m \leqq n(s-r). \tag{1. 3}$$

Here $\operatorname{rank}_{C(\lambda)}$ denotes the rank over the field $\mathbb{C}(\lambda)$ of rational functions with complex coefficients.

Unfortunately, the proof in [5] does not indicate how a meromorphic transformation could be constructed if the criterion is satisfied.

In the case $r=0$, the above theorem is due to Wagenführer [8]. He even constructs a fundamental solution

$$Y(z)=\sum_{r=0}^{\infty} Y_r z\cdot z^J$$

of (1. 1) from properties of $\mathbf{A}_m^{(0)}(\lambda)$, if the criterion is satisfied for $r=0$.

Using similar properties of $\mathbf{A}_m^{(r)}(\lambda)$ we want to construct a meromorphic transformation that reduces the Poincaré rank to r if criterion (1. 3) is satisfied. This can be regarded as an answer to the third question. We cannot expect to construct a (formal) fundamental solution, if criterion (1. 3) is fulfilled for $r>0$, because nothing can be said about possible rational powers of z.

As in [5] the main difficulty in the proof consists in showing that certain properties of $\mathbf{A}_m^{(r)}(\lambda)$ do not change if A_s is replaced by $A_s-\alpha I$, $\alpha\in\mathbb{C}$. In the case $r=0$ this problem does not arise.

As another main result for arbitrary systems (1. 1) we obtain some formal invariant quantities that are defined in terms of the power series coefficients of $A(z)$ but that are also related to a formal fundamental solution of (1. 1) in a simple manner.

We note that the above problems do not depend upon the convergence of the power series, one may as well assume that they are formal power series. Then the transformations (1. 2) are called *formal meromorphic* or *formal analytic*, resp..

After completion of this work E. Wagenführer [9] communicated to us that he found a quite different way to construct a meromorphic rank reducing transformation.

2. The formal meromorphic invariants $\chi^{(r)}(\lambda)$

First we consider arbitrary polynomial matrices $A(\lambda)$, i.e. matrices of polynomials in λ.

Definition 2. 1. For a polynomial matrix $A(\lambda)$ of rank k (over $\mathbb{C}(\lambda)$) let the "characterizing polynomial" $\chi_A(\lambda)$ denote the greatest common divisor of all minors of $A(\lambda)$ of order k normalized to leading coefficient one.

In [3], vol. I, p. 139f. it is shown that the characterizing polynomial is invariant with respect to elementary operations, i.e. multiplication of a row or column by a constant $\neq 0$, interchange of any two rows or columns and addition of a polynomial multiple of one row (column) to another row (column).

For a polynomial matrix $A(\lambda)$ of n rows and columns we define the defect

$$\operatorname{def} A(\lambda)=n-\operatorname{rank}_{C(\lambda)} A(\lambda)$$

i.e. the dimension of the kernel of $A(\lambda)$.

The following lemma is useful in this section. It relates the characterizing polynomials of certain polynomial matrices and their submatrices. To simplify notation we sometimes omit the argument λ of polynomial matrices.

Lemma 2. 1. *Let $B(\lambda)$ be a polynomial matrix with block structure*

$$B=\begin{pmatrix} B_{11} & 0 & 0 \\ B_{21} & B_{22} & 0 \\ B_{31} & B_{32} & B_{33} \end{pmatrix},$$

where $B_{ij}(\lambda)$ are n_i by n_j matrices. We assume that

$$\operatorname{def} B(\lambda)=\operatorname{def} B_{22}(\lambda)=d. \tag{2. 1}$$

Let $E(\lambda)$ and $F(\lambda)$ denote the block matrices

$$E=\begin{pmatrix} B_{11} & 0 \\ B_{12} & B_{22} \end{pmatrix}, \qquad F=\begin{pmatrix} B_{22} & 0 \\ B_{32} & B_{33} \end{pmatrix}.$$

Then $\chi_E(\lambda)$ divides $\chi_B(\lambda)$, $\chi_{B_{22}}(\lambda)$ divides $\chi_F(\lambda)$ and

$$\frac{\chi_B(\lambda)}{\chi_E(\lambda)}=\frac{\chi_F(\lambda)}{\chi_{B_{22}}(\lambda)}.$$

Proof. Obviously

$$\operatorname{def} B_{22}(\lambda)\leqq \operatorname{def} E(\lambda)\leqq \operatorname{def} B(\lambda)$$

and

$$\operatorname{def} B_{22}(\lambda)\leqq \operatorname{def} F(\lambda)\leqq \operatorname{def} B(\lambda).$$

Then (2. 1) implies that the defects of B_{22}, F, E and B are equal to d. Since $\operatorname{def} B_{22}(\lambda)=d$ there are elementary operations on the rows and columns n_1+1 to n_1+n_2 that transform $B(\lambda)$ to a block structure

$$\begin{pmatrix} B_{11} & 0 & 0 & 0 \\ \tilde{B}_{21} & \tilde{B}_{22} & 0 & 0 \\ \tilde{\tilde{B}}_{21} & 0 & 0 & 0 \\ B_{31} & \tilde{B}_{32} & \tilde{\tilde{B}}_{32} & B_{33} \end{pmatrix} \tag{2. 2}$$

where the four block rows and block columns contain n_1, n_2-d, d and n_3 rows and columns, respectively (see [3], ch. VI, § 2). Using elementary operations on the last n_3+d columns $\tilde{\tilde{B}}_{32}=0$ can be attained. Hereby B_{33} is changed to some $\tilde{B}_{33}$.

Then we have

$$\chi_{B_{22}}(\lambda)=\det \tilde{B}_{22}(\lambda) \tag{2. 3}$$

and

$$\chi_F(\lambda)=\det \tilde{B}_{22}(\lambda)\cdot \det \tilde{B}_{33}(\lambda), \tag{2. 4}$$

where we assume that $\det \tilde{B}_{22}(\lambda)$ and $\det \tilde{B}_{33}(\lambda)$ are normalized.

Again because of [3], ch. VI, § 2, th. 2 there are elementary operations on the first n_1+n_2 rows of (2. 2) (with $\tilde{B}_{32}=0$) transforming it to

$$\begin{pmatrix} \tilde{B}_{11} & 0 & 0 \\ 0 & 0 & 0 \\ \tilde{B}_{31} & 0 & \tilde{B}_{33} \end{pmatrix}$$

where the block rows (columns) contain n_1+n_2-d, d and n_3 rows (columns), respectively. Then we have

$$\chi_B(\lambda) = \det \tilde{B}_{11}(\lambda) \cdot \det \tilde{B}_{33}(\lambda) \tag{2. 5}$$

and

$$\chi_E(\lambda) = \det \tilde{B}_{11}(\lambda). \tag{2. 6}$$

where we assume again that $\det \tilde{B}_{11}(\lambda)$ is normalized. The formulae (2. 3)–(2. 6) prove the lemma.

Now we fix a positive integer $r \leqq s-1$ and consider the matrices $\mathbf{A}_m(\lambda) = \mathbf{A}_m^{(r)}(\lambda)$ of the first section. Since r is fixed up to the end of the proof of theorem 2. 4 we omit the superscript r of $\mathbf{A}_m^{(r)}(\lambda)$. Further let $\chi_m(\lambda) = \chi_m^{(r)}(\lambda) = \chi_{A_m}(\lambda)$ denote the characterizing polynomials of $\mathbf{A}_m(\lambda)$.

We apply lemma 2. 1 to $B(\lambda) = \mathbf{A}_{m+2}(\lambda)$ and $n_2 = mn$, $n_1 = n_3 = n$. Then with the notations of the lemma we have

$$\begin{aligned} E(\lambda) &= \mathbf{A}_{m+1}(\lambda), \\ F(\lambda) &= \mathbf{A}_{m+1}(\lambda) - \mathbf{H}_{m+1}, \\ B_{22}(\lambda) &= \mathbf{A}_m(\lambda) - \mathbf{H}_m, \end{aligned}$$

where $\mathbf{H}_l$ denotes the ln by ln matrix containing one's in the sn^{th} subdiagonal and zeros everywhere else.

In order to verify the assumptions of the lemma we need two statements of [5] that are proved independently in the third section, too.

By [5], lemma 3. 1

$$\operatorname{def} \mathbf{A}_l(\lambda) = \operatorname{def}(\mathbf{A}_l(\lambda) - \alpha \mathbf{H}_l) \tag{2. 7}$$

for all complex α and $l = s, s+1, \ldots$. By [5], lemma 3. 2 there exists $N \leqq n(s-r)$ such that for all $l \geqq N$

$$\operatorname{def} \mathbf{A}_l(\lambda) = \operatorname{def} \mathbf{A}_N(\lambda). \tag{2. 8}$$

Because of (2. 7) and (2. 8) condition (2. 1) is satisfied for $m \geqq N$. In the next section (corollary (3. 14)) we show that not only the defects but also the characterizing polynomials of $\mathbf{A}_l(\lambda)$ and $\mathbf{A}_l(\lambda) - \alpha \mathbf{H}_l$ agree. Thus lemma 2. 1 yields

Theorem 2. 1. *The characterizing polynomials $\chi_m(\lambda)$ divide $\chi_{m+1}(\lambda)$ for $m \geqq N$ and the quotient $\dfrac{\chi_{m+1}(\lambda)}{\chi_m(\lambda)}$ is independent of $m \geqq N$.*

Thus we attach the polynomial

$$\chi^{(r)}(\lambda) = \frac{\chi_{m+1}(\lambda)}{\chi_m(\lambda)} \quad (m \geqq N) \tag{2. 9}$$

to the system (1. 1) of differential equations. It is easy to see that the polynomials $\chi_m(\lambda)$ remain unchanged under analytic transformations (cf. [5], proof of lemma 3. 1 (ii)). But simple examples show that $\chi_m(\lambda)$ are not formal meromorphic invariants. In contrast to this behavior the following theorem asserts that the polynomials $\chi^{(r)}(\lambda)$ are not changed by formal meromorphic transformations.

Theorem 2. 2. *The polynomials* $\chi^{(r)}(\lambda)$, $r=1,\ldots,s-1$ *are formal meromorphic invariants of* (1. 1).

Proof. Any formal meromorphic transformation can be decomposed into formal analytic transformations and transformations of the form $y=z^K\tilde{y}$, where K is a diagonal matrix whose first v diagonal elements are one and the remaining ones are zero (cf. [6], p. 384ff.). Like $\chi_m(\lambda)$ the polynomials $\chi^{(r)}(\lambda)$ are not changed by formal analytic transformations. Hence it suffices to show that $\chi^{(r)}(\lambda)$ is invariant with respect to a transformation $y=z^K\tilde{y}$ of the above type. Such a transformation takes (1. 1) into

$$z^{s+1}\tilde{y}' = \tilde{A}(z)\tilde{y} \quad \text{with} \quad \tilde{A}(z) = z^{-K}A(z)z^K - z^s K.$$

If we write

$$A(z) = \begin{pmatrix} A^1(z) & A^2(z) \\ A^3(z) & A^4(z) \end{pmatrix},$$

where $A^1(z)$ is a v by v matrix, then

$$\tilde{A}(z) = \begin{pmatrix} A^1(z) - z^s I, & \frac{1}{z}A^2(z) \\ zA^3(z), & A^4(z) \end{pmatrix}.$$

Since $\chi^{(r)}(\lambda)$ is not changed by multiplying both sides of (1. 1) by z, we can assume that $A^2(0)=0$ and hence $\tilde{A}(z)$ is a power series in z.

Now construct $\tilde{\mathbf{A}}_m(\lambda)$ from $\tilde{A}(z)$ in the same way as $\mathbf{A}_m(\lambda)$ was built from $A(z)$. We note that $\tilde{\mathbf{A}}_m(\lambda)$ can be obtained from $\mathbf{A}_{m+1}(\lambda)$ in the following way. First delete the first v rows and columns of $\mathbf{A}_{m+1}(\lambda)$. In the resulting matrix $\bar{\mathbf{A}}_{m+1}(\lambda)$ delete the last $n-v$ rows and columns. The matrix $\bar{\bar{\mathbf{A}}}_{m+1}(\lambda)$ obtained in this way agrees with $\tilde{\mathbf{A}}_m(\lambda)$ except for permutations of rows and columns. This implies

$$\bar{\bar{\chi}}_{m+1}(\lambda) = \tilde{\chi}_m(\lambda), \quad m=0,1,2,\ldots \tag{2. 10}$$

where here and in the sequel $\bar{\chi}_m(\lambda)$, $\bar{\bar{\chi}}_m(\lambda)$, $\tilde{\chi}_m(\lambda)$ denote the characterizing polynomials of $\bar{\mathbf{A}}_m(\lambda)$, $\bar{\bar{\mathbf{A}}}_m(\lambda)$, $\tilde{\mathbf{A}}_m(\lambda)$, respectively.

Now we apply lemma 2.1 to $B(\lambda)=\mathbf{A}_{m+2}(\lambda)$ with $n_1=\nu$, $n_2=(m+1)n-\nu$ and $n_3=n$. Because of $A^2(0)=0$ the blocks of $B(\lambda)$ above the diagonal blocks vanish as required in lemma 2.1. Further

$$\begin{aligned} E(\lambda) &= \mathbf{A}_{m+1}(\lambda), \\ F(\lambda) &= \bar{\mathbf{A}}_{m+2}(\lambda), \\ B_{22}(\lambda) &= \bar{\mathbf{A}}_{m+1}(\lambda) \end{aligned}$$

are the matrices built in lemma 2.1. Since obviously

$$\operatorname{def}(\mathbf{A}_m(\lambda)-\mathbf{H}_m) \leqq \operatorname{def}(\bar{\mathbf{A}}_{m+1}(\lambda)) \leqq \operatorname{def}(\mathbf{A}_{m+1}(\lambda)),$$

the properties (2.7) and (2.8) imply that the hypothesis (2.1) of lemma 2.1 is satisfied. Hence

$$\frac{\chi_{m+2}(\lambda)}{\chi_{m+1}(\lambda)} = \frac{\bar{\chi}_{m+2}(\lambda)}{\bar{\chi}_{m+1}(\lambda)} \quad (m \geqq N). \tag{2.11}$$

Now we use lemma 2.1 once more and apply it to $B(\lambda)=\bar{\mathbf{A}}_{m+2}(\lambda)$ with $n_1=n$, $n_2=mn$ and $n_3=n-\nu$. Then

$$\begin{aligned} E(\lambda) &= \bar{\bar{\mathbf{A}}}_{m+2}(\lambda), \\ F(\lambda) &= \bar{\mathbf{A}}_{m+1}(\lambda)-\bar{\mathbf{H}}_{m+1}, \\ B_{22}(\lambda) &= \bar{\bar{\mathbf{A}}}_{m+1}(\lambda)-\mathbf{H}_m. \end{aligned}$$

are the corresponding matrices of lemma 2.1. As before

$$\operatorname{def}(\mathbf{A}_{m-1}(\lambda)-2\mathbf{H}_{m-1}) \leqq \operatorname{def}(\bar{\bar{\mathbf{A}}}_{m+1}(\lambda)-\mathbf{H}_m) \leqq \operatorname{def}\bar{\mathbf{A}}_{m+1}(\lambda) \leqq \operatorname{def}\mathbf{A}_{m+1}(\lambda)$$

and because of (2.7) and (2.8) hypothesis (2.1) of lemma 2.1 is satisfied for $m \geqq N+1$. In the third section (corollary 1) we show that the characterizing polynomials of $\tilde{\mathbf{A}}_l(\lambda)-\alpha\mathbf{H}_l$ and of $\bar{\mathbf{A}}_l(\lambda)-\alpha\bar{\mathbf{H}}_l$ do not depend upon α. Because of (2.10) this is also true for $\bar{\bar{\mathbf{A}}}_m(\lambda)-\alpha\mathbf{H}_m$.

Thus lemma 2.1 yields

$$\frac{\bar{\chi}_{m+2}(\lambda)}{\bar{\bar{\chi}}_{m+2}(\lambda)} = \frac{\bar{\chi}_{m+1}(\lambda)}{\bar{\bar{\chi}}_{m+1}(\lambda)} \tag{2.12}$$

The theorem follows easily from (2.10)—(2.12).

The system (1.1) of differential equations has a formal fundamental solution matrix (see [1]) which can be expressed in the form

$$Y(z) = F(z^{\frac{1}{p}})\, z^J \exp[Q(z^{-\frac{1}{p}})], \tag{2.13}$$

where p is a positive integer, $Q(u)=\operatorname{diag}\{q_1(u),\ldots,q_n(u)\}$ is a diagonal matrix of polynomials without constant term, J is a constant matrix commuting with $Q(u)$ for all complex u and $F(z^{\frac{1}{p}})$ is a formal meromorphic series in $z^{\frac{1}{p}}$ with $\det F(z^{\frac{1}{p}}) \not\equiv 0$. In the remainder of this section we want to find a relation between the polynomials $\chi^{(r)}(\lambda)$ corresponding to (1.1) and the polynomials $q_j(u)$. Since rational powers of z appear in (2.13) we have to investigate the behavior of the polynomials $\chi^{(r)}(\lambda)$ when $z=t^p$, p a positive integer, is substituted in (1.1). From now on we will indicate the dependence of $\mathbf{A}_m^{(r)}(\lambda)$, $\chi_m^{(r)}(\lambda)$ of r again.

The substitution $z=t^p$, $y(z)=x(t)$ with a positive integer p transforms (1. 1) to

$$t^{sp+1}\frac{d}{dt}x=pA(t^p)x. \tag{2. 14}$$

Now let $\tilde{\mathbf{A}}_m^{(r)}(\lambda)$ be constructed from $\tilde{A}(t)=pA(t^p)$ in the same way as $\mathbf{A}_m^{(r)}(\lambda)$ from $A(z)$. Then the matrix $\tilde{\mathbf{A}}_{pm}^{(pr)}(p\lambda)$ can be transformed to a block diagonal form

$$\operatorname{diag}\left(p\mathbf{A}_m^{(r)}(\lambda),\, p\mathbf{A}_m^{(r)}(\lambda)-\mathbf{H}_m,\ldots,\, p\mathbf{A}_m^{(r)}(\lambda)-(p-1)\,\mathbf{H}_m\right)$$

by permutations of rows and columns.

Since the characterizing polynomials of $\mathbf{A}_m^{(r)}-\alpha\mathbf{H}_m$ do not depend upon α the characterizing polynomial of $\tilde{\mathbf{A}}_{pm}^{(pr)}(p\lambda)$ equals $[\chi_m^{(r)}(\lambda)]^p$. The definition of the polynomials attached to a differential equation immediately yields

$$\tilde{\chi}^{(pr)}(p\lambda)=\chi^{(r)}(\lambda), \tag{2. 15}$$

where $\tilde{\chi}^{(r)}(\lambda)$ denotes the r^{th} polynomials attached to (2. 14).

Theorem 2. 3. *Let the system* (1. 1) *of differential equations have a formal fundamental solution* (2. 13). *Let r be a positive integer and let $p\cdot\partial q_j$ denote the degree of the polynomial $q_j(u)$. Then the degree of the polynomial $\chi^{(r)}(\lambda)$ attached to* (1. 1) *in theorem* (2. 1) *is equal to the number of j such that $\partial q_j\leqq r$.*

The zeros of $\chi^{(r)}(\lambda)$, according to their multiplicity, are the coefficients of z^{-r-1} of those functions $\frac{d}{dz}\left(q_j(z^{-\frac{1}{p}})\right)$ such that $\partial q_j\leqq r$.

Proof. Because of a possible substitution $z=t^p$ and (2. 15) we can assume that $p=1$. Then the formal meromorphic transformation $y=F(z)\tilde{y}$ leads from (1. 1) to a differential equation in $\mathbb{C}^n$ with formal fundamental solution $z^J\exp\left(Q(z^{-1})\right)$, i. e.

$$z^{s+1}\tilde{y}'=\left[z^{s+1}\frac{d}{dz}Q(z^{-1})+z^sJ\right]\tilde{y}, \tag{2. 16}$$

since J and Q commute. Because of the meromorphic invariance of $\chi^{(r)}(\lambda)$ we can assume that (1. 1) has already form (2. 16). Because Q and J commute the coefficient matrix of (2. 16) is block diagonal (see [3], vol. I, p. 223).

Now the definition of the polynomials attached to a differential equation shows that if $A(z)=\operatorname{diag}\left(A_1(z),A_2(z)\right)$ is block diagonal then $\chi^{(r)}(\lambda)$ is the product of the polynomials $\chi_i^{(r)}(\lambda)$, $i=1,2$ attached to the systems $z^{s+1}y'=A_i(z)\,y$, $i=1,2$. Hence it is sufficient to prove the theorem for $Q(z^{-1})=q(z^{-1})I$ in (2. 16).

If we write

$$\frac{d}{dz}\left(q(z^{-1})\right)=\sum_{\nu=1}^{s}a_\nu z^{-\nu-1}$$

then the coefficient matrix of (2. 16) has the form $\sum_{\nu=1}^{s}a_\nu z^{s-\nu}I+z^sJ$. Thus if the degree of q exceeds r at least one of the coefficients $a_{r+1},\ldots,a_s$ does not vanish and the definition immediately shows that $\chi_m^{(r)}(\lambda)=1$ for all m. Hence $\chi^{(r)}(\lambda)=1$ and the theorem is proved in this case. If the degree of q is less or equal r then we have $a_{r+1}=\cdots=a_s=0$ and this yields $\chi_m^{(r)}(\lambda)=(\lambda-a_r)^{(m-s+r)n}$ for all $m>s-r$. This implies $\chi^{(r)}(\lambda)=(\lambda-a_r)^n$ and the theorem is proved in this case, too.

One can define polynomials $\chi^{(r)}(\lambda)$ attached to (1. 1) also for rational r with $0<r<s$. To that purpose choose any p such that pr is an integer and define $\chi^{(r)}(\lambda)$ to be the left hand side of (2. 15). For these polynomials theorem 2. 3 is valid too.

In [5], def. 3. 2 (meromorphically) invariant numbers $i^{(r)}$ were attributed to system (1. 1). These were defined by

$$i^{(r)} := n(s-r) - \operatorname{def} \mathbf{A}_m^{(r)}(\lambda)$$

for m sufficiently large. The following relation between $i^{(r)}$ and the formal fundamental solution (2. 13) was stated ([5], thm. 3. 2):

$$i^{(r)} = \sum_{j=1}^{n} \max(\partial q_j - r, 0). \tag{2. 17}$$

Comparison of (2. 17) and 2. 3 yields

Corollary. *The invariant number $i^{(r)}$ of a system* (1. 1) *vanishes if and only if the polynomial $\chi^{(r)}(\lambda)$ attached to* (1. 1) *has degree n.*

Finally we remark that theorem 2. 3 gives a characterization of the number of j, such that $\partial q_j > r$, that is simpler than corollary 2 in [5], section 3. There it was proved that this number is given by

$$\frac{[i^{(r)} - i^{(r+\delta)}]}{\delta}$$

for $\delta > 0$ sufficiently small. Because of theorem 2. 3 this number is equal to

$$n - \operatorname{degree} \chi^{(r)}(\lambda).$$

This is simpler than the above formula but the computation of $\chi^{(r)}(\lambda)$ is more involved than that of the defects $d_m^{(r)}$ of $\mathbf{A}_m^{(r)}(\lambda)$.

3. Construction of a transformation reducing the Poincaré rank.

Consider the system (1. 1) of differential equations. Fix a positive integer $r<s$ and let

$$Dy := -z^{s+1} \frac{d}{dz} y + A(z)\, y - \lambda z^{s-r} y \quad (y \in \mathfrak{O}) \tag{3. 1}$$

denote a differential operator corresponding to (1. 1) defined on $\mathfrak{O} = (\mathbb{C}[\lambda]\,[[z]])^n$, i. e. the set of n-tuples of (formal) power series in z whose coefficients are polynomials in a second variable λ. The set $\mathfrak{O}$ is a vector space over $\mathbb{C}$ and a module over the ring $\mathbb{C}[\lambda]$, D is a $\mathbb{C}[\lambda]$-linear operator on $\mathfrak{O}$. The matrices $\mathbf{A}_m(\lambda)$ —and also $\bar{\mathbf{A}}_m(\lambda)$—in section 2 are closely related to D in the following way.

If we expand the D-images of $z^k e_j$, e_j denoting the j-th unit vector,

$$D(z^k e_j) = \sum_{l=0}^{\infty} \sum_{\nu=1}^{n} a_{jk}^{\nu l}(\lambda)\, z^l\, e_\nu \tag{3. 2}$$

for $k = 0, 1, \ldots$ and $j = 1, \ldots, n$ with scalar polynomials $a_{jk}^{\nu l}(\lambda)$ (at most linear) in λ, then $\mathbf{A}_m(\lambda)$ is the matrix whose element in the $(kn+j)^{\text{th}}$ column and $(ln+\nu)^{\text{th}}$ row is $a_{jk}^{\nu l}(\lambda)$.

Further let d_m denote the defect and $\chi_m(\lambda)$ denote the characterizing polynomial of $\mathbf{A}_m(\lambda)$ (see section 2). Finally for positive integers k and m we define $\mathfrak{D}_{k,m} \subset \mathfrak{D}$ as the set of polynomials in λ and z of λ-degree $< k$ and z-degree $< m$ with coefficients in $\mathbb{C}^n$.

Lemma 3. 1. *Let μ be a complex number and k, m positive integers. Consider the $\mathbb{C}$-vectorspace $\mathfrak{K}(\mu, k, m)$ of all $x \in \mathfrak{D}_{k,m}$ such that*

$$Dx \in (\lambda-\mu)^k \mathfrak{D} + z^m \mathfrak{D}. \tag{3. 3}$$

Then the dimension of $\mathfrak{K}(\mu, k, m)$ is $v(\mu, \chi_m)+k d_m$ for large k, where $v(\mu, \chi_m)$ denotes the multiplicity of μ as a zero of $\chi_m(\lambda)$.

(3. 3) simply means that in the expansion

$$Dx = \sum_{j=0}^{\infty} \left(\sum_{l \geq 0} c_{jl} (\lambda-\mu)^l \right) z^j$$

the coefficients $c_{jl} \in \mathbb{C}^n$ vanish if $j < m$ and $l < k$. Thus $\mathfrak{K}(\mu, k, m)$ is the kernel of the linear operator mapping $x \in \mathfrak{D}_{k,m}$ to the vector $(c_{j,l})_{j=0,\ldots,m-1}^{l=0,\ldots,k-1}$ in $\mathbb{C}^{mnk}$.

This lemma can be seen as an application of theorem 1. 19 in [8] (there matrix notation is used), but we will prove it independently in a way that one sees how a basis of $\mathfrak{K}(\mu, k, m)$ can be constructed.

Proof. According to [3], VI, § 2 there are polynomial matrices $P(\lambda)$ and $Q(\lambda)$ in λ with constant nonvanishing determinant such that

$$P(\lambda)^{-1} \mathbf{A}_m(\lambda) Q(\lambda) = S(\lambda) \tag{3. 4}$$

is the Smith normal form of $\mathbf{A}_m(\lambda)$, i.e. a diagonal matrix with diagonal entries $i_j(\lambda)$, $j=1,\ldots,nm-d_m$ (the invariant polynomials of $\mathbf{A}_m(\lambda)$) and d_m zeros. The $i_j(\lambda)$ are normalized to leading coefficient one and $i_{j-1}(\lambda)$ divides $i_j(\lambda)$ for $j>1$. $\chi_m(\lambda)$ is the product of the invariant polynomials $i_j(\lambda)$.

Now we construct polynomials $p_j, q_j \in \bigcup_k \mathfrak{D}_{k,m}$ from the columns of $P(\lambda)$, $Q(\lambda)$ respectively by forming, e.g.,

$$p_j = \sum_{\nu=1}^{n} \sum_{k=0}^{m-1} p_{kn+\nu,j}(\lambda)\, z^k\, e_\nu$$

for a column $(p_{ij}(\lambda))_{i=1,\ldots,nm}$ of $P(\lambda)$.

Then because of (3. 2) and (3. 4)

$$\begin{aligned} Dq_j &\equiv i_j(\lambda)\, p_j \bmod z^m \mathfrak{D} \quad (j=1,\ldots,nm-d_m), \\ Dq_j &\equiv 0 \qquad \bmod z^m \mathfrak{D} \quad (j=nm-d_m+1,\ldots,nm) \end{aligned} \tag{3. 5}$$

and

$$\{q_j|_{\lambda=\alpha},\ j=1,\ldots,nm\} \quad \text{and} \quad \{p_j|_{\lambda=\alpha},\ j=1,\ldots,nm\} \tag{3. 6}$$

are linearly independent for all $\alpha \in \mathbb{C}$, where

$$x|_{\lambda=\alpha} = \sum_{\nu=0}^{\infty} \left(\sum_l x_{\nu l} \alpha^l\right) z^\nu \quad \text{for} \quad x = \sum_{\nu=0}^{\infty} \left(\sum_l x_{\nu l} \lambda^l\right) z^\nu \in \mathfrak{D}.$$

Now we define $q_{jv} \in \mathfrak{O}_{k,m}$ by

$$q_{jv} \equiv (\lambda-\mu)^v q_j \bmod (\lambda-\mu)^k \mathfrak{O} \quad (j=1,\ldots,nm,\ v=0,\ldots,k-1).$$

Because of (3. 6) they form a $\mathbb{C}$-basis of $\mathfrak{O}_{k,m}$. Now if $q=\sum_{j,v} \alpha_{jv} q_{jv}$ is an arbitrary element of $\mathfrak{O}_{k,m}$ then because of (3. 5)

$$Dq \equiv \sum_{v=0}^{k-1} \sum_{j=1}^{nm-d_m} \alpha_{jv}(\lambda-\mu)^v i_j(\lambda)\, p_j \bmod (\lambda-\mu)^k \mathfrak{O} + z^m \mathfrak{O}.$$

Because $p_j|_{\lambda=\mu}$, $j=1,\ldots,nm$ are linearly independent, q is in $\mathfrak{R}(\mu,k,m)$ if and only if $\alpha_{jv}=0$ when $j \leqq nm-d_m$ and $v+v(\mu,i_j) \leqq k-1$. Therefore

$$\dim_{\mathbb{C}} \mathfrak{R}(\mu,k,m) = \sum_{j=1}^{nm-d_m} \min\big(k, v(\mu,i_j)\big) + k d_m \tag{3. 7}$$

and the lemma is proved.

Besides $\mathfrak{R}(\mu,k,m)$ we introduce two other $\mathbb{C}$-vector spaces.

Definition 3. 1. For positive integers k and m let $\mathfrak{C}(k,m)$ denote the set of all $x \in \mathfrak{O}_{k,m}$ such that $Dx \equiv 0 \bmod z^m$. For complex μ let $\mathfrak{I}(\mu,k,m)$ denote the set of $y \in \mathfrak{O}_{1,m}$ such that there exists an $x \in \mathfrak{O}_{k,m}$ such that

$$Dx \equiv (\lambda-\mu)^k y \bmod z^m \mathfrak{O}.$$

Then $\mathfrak{C}(k,m)$ is the intersection of all $\mathfrak{R}(\mu,k,m)$, $\mu \in \mathbb{C}$, i.e. it contains the elements common to all the "kernels" $\mathfrak{R}(\mu,k,m)$.

Now let φ map $x \in \mathfrak{R}(\mu,k,m)$ to the uniquely determined $y \in \mathfrak{O}_{1,m}$ such that

$$Dx \equiv (\lambda-\mu)^k y \bmod z^m \mathfrak{O}.$$

This mapping is well defined because D increases the λ-degree at most by one. Then $\mathfrak{I}(\mu,k,m)$ is the image of φ and $\mathfrak{C}(k,m)$ is its kernel.

Hence

$$\dim \mathfrak{R}(\mu,k,m) = \dim \mathfrak{C}(k,m) + \dim \mathfrak{I}(\mu,k,m). \tag{3. 8}$$

The importance of these new vector spaces is revealed by

Theorem 3. 1. *Let α be a complex number and $\tilde{r}$ be a positive integer. Then for all positive integers k and m the dimension of $\mathfrak{C}(k,m)$ is not changed by replacing D by $D-\alpha z^{s-r+\tilde{r}} \mathrm{id}_{\mathfrak{O}}$. Further for all positive integers m and complex μ there exist k_0 such that $\mathfrak{I}(\mu,k,m)$ is not changed by this substitution for $k \geqq k_0$.*

Proof. As in [5], section 3 we define scalar polynomials $p_{ij}(t,z)$, $i,j=0,1,\ldots$ of two variables t and z by the following recursion (arguments are omitted):

$$p_{i0}=1,\ p_{0j+1}=z^{r+1}\frac{d}{dz}p_{0j}+tz^{\tilde{r}}p_{0j} \quad (i,j=0,1,\ldots)$$

$$p_{ij+1}=z^{r+1}\frac{d}{dz}p_{ij}+tz^{\tilde{r}}p_{ij}+p_{i-1,j+1} \quad (i=1,2,\ldots;\ j=0,1,\ldots). \tag{3. 9}$$

Now choose any complex α and any positive integer k. Let $\tilde{\mathfrak{C}}(k, m)$ denote the vector space of definition 3.1 formed for $D-\alpha z^{s-r+\tilde{r}}\,\mathrm{id}_{\mathfrak{O}}$ instead of D. We have to show that the dimensions of $\mathfrak{C}(k, m)$ and $\tilde{\mathfrak{C}}(k, m)$ are equal.

To that purpose define a mapping

$$f^{\alpha}: \begin{cases} \mathfrak{O}_{k,m} & \longrightarrow \mathfrak{O}_{k,m}\,, \\ \sum_{i=0}^{k-1} \lambda^i v_i & \longrightarrow \sum_{i=0}^{k-1} \lambda^i \tilde{v}_i \end{cases} \tag{3.10}$$

where $\tilde{v}_0, \ldots, \tilde{v}_{k-1} \in \mathfrak{O}_{1,m}$ are formed from $v_0, \ldots, v_{k-1} \in \mathfrak{O}_{1,m}$ according to

$$\tilde{v}_i \equiv \sum_{j=0}^{k-1-i} p_{ij}(\alpha, z)\, v_{i+j} \bmod z^m \mathbb{C}^n[z]. \tag{3.11}$$

First we show that f^{α} maps $\mathfrak{C}(k, m)$ into $\tilde{\mathfrak{C}}(k, m)$. Note that the congruence $D(\Sigma \lambda^i v_i) \equiv 0 \bmod z^m \mathfrak{O}$ is equivalent to $k+1$ congruences $\bmod z^m$

$$-z^{s+1}\frac{d}{dz} v_i + A(z)\, v_i \equiv z^{s-r}\, v_{i-1} \quad (i=1, \ldots, k-1)\,,$$

$$-z^{s+1}\frac{d}{dz} v_0 + A(z)\, v_0 \equiv z^{s-r} v_{k-1} \equiv 0\,.$$

A straight forward computation using these relations and (3.9), (3.11) shows that under the assumption $D(\Sigma \lambda^i v_i) \equiv 0 \bmod z^m \mathfrak{O}$ we have

$$-z^{s+1}\frac{d}{dz} \tilde{v}_i + A(z)\, \tilde{v}_i \equiv z^{s-r} \tilde{v}_{i-1} + \alpha z^{s-r+\tilde{r}} \tilde{v}_i \quad (i=1, \ldots, k-1)\,,$$

$$-z^{s+1}\frac{d}{dz} \tilde{v}_0 + A(z)\, \tilde{v}_0 - \alpha z^{s-r+\tilde{r}}\, \tilde{v}_0 \equiv z^{s-r}\, \tilde{v}_{k-1} \equiv 0$$

$\bmod z^m$ and this implies

$$(D-\alpha z^{s-r+\tilde{r}}\,\mathrm{id}_{\mathfrak{O}}) \left(\sum_{i=0}^{k-1} \lambda^i \tilde{v}_i \right) \equiv 0 \bmod z^m \mathfrak{O}$$

and thus $\Sigma \lambda^i \tilde{v}_i = f^{\alpha}(\Sigma \lambda^i v_i)$ is in $\tilde{\mathfrak{C}}(k, m)$ whenever $\Sigma \lambda_i v^i \in \mathfrak{C}(k, m)$. Essentially the same computation shows that $f^{-\alpha}$ maps $\tilde{\mathfrak{C}}(k, m)$ into $\mathfrak{C}(k, m)$. Further f^{α} and $f^{-\alpha}$ are $\mathbb{C}$-linear and one-to-one because of $p_{i0} = 1$, $i = 0, 1, \ldots$, and (3.11). This yields the first part of the theorem.

To prove the second part we first show that

$$\mathfrak{J}(\mu, k, m) \subseteqq \tilde{\mathfrak{J}}(\mu, k+m, m) \quad \text{for all} \quad k \in \mathbb{N}, \tag{3.12}$$

where $\tilde{\mathfrak{J}}(\mu, k, m)$ denotes the set of definition 3.1 formed for $D-\alpha z^{s-r+\tilde{r}}\,\mathrm{id}_{\mathfrak{O}}$ instead of D.

So choose any $y \in \mathfrak{J}(\mu, k, m)$ and a corresponding $x = \sum_{i=0}^{k-1} (\lambda-\mu)^i v_i$ in $\mathfrak{O}_{k,m}$ with coefficients $v_i \in \mathfrak{O}_{1,m}$ such that

$$Dx \equiv (\lambda-\mu)^k y \bmod z^m \mathfrak{O}\,.$$

Then we have the following congruences $\bmod z^m$

$$-z^{s+1}\frac{d}{dz}v_i + A(z)\,v_i - \mu z^{s-r}v_i \equiv z^{s-r}v_{i-1} \quad (i=1,\ldots,k-1),$$

$$-z^{s+1}\frac{d}{dz}v_0 + A(z)\,v_0 - \mu z^{s-r}v_0 \equiv 0,$$

$$-z^{s-r}v_{k-1} \equiv y.$$

Now we define $\bar{v}_0,\ldots,\bar{v}_{k+m-1} \in \mathfrak{O}_{1,m}$ similar to (3. 11) by

$$\bar{v}_i \equiv \sum_{\substack{l=0 \\ l \geqq i-m}}^{k-1} (-1)^{l+m+i}\, p_{k-1-l,l+m-i}(\alpha, z)\, v_l \bmod z^m.$$

Since z^ν divides $p_{i\nu}(t, z)$ for all i, ν we first conclude that $\bar{v}_0 = 0$. A straight forward computation analogous to the previous one shows that

$$-z^{s+1}\frac{d}{dz}\bar{v}_i + A(z)\,\bar{v}_i - \mu z^{s-r}\bar{v}_i \equiv \alpha z^{s-r+\tilde{r}}\bar{v}_i + z^{s-r}\bar{v}_{i-1} \bmod z^m$$

for $i=1,\ldots,k+m-1$ and that $y \equiv -z^{s-r}\bar{v}_{k+m-1} \bmod z^m$. This implies

$$(D-\alpha z^{s-r+\tilde{r}}\,\mathrm{id}_{\mathfrak{O}})\,(\Sigma(\lambda-\mu)^i\bar{v}_i) \equiv (\lambda-\mu)^{k+m}y \bmod z^m\mathfrak{O}$$

and thus $y \in \tilde{\mathfrak{J}}(\mu, k+m, m)$. The proof of (3. 12) is now complete.

If we interchange the roles of D and $D-\alpha z^{s-r+\tilde{r}}$, we obtain

$$\tilde{\mathfrak{J}}(\mu, k, m) \subseteqq \mathfrak{J}(\mu, k+m, m) \tag{3. 13}$$

for all positive integers k and m. Since

$$\mathfrak{J}(\mu, k, m) \subseteqq \mathfrak{J}(\mu, k+1, m) \subseteqq \mathfrak{O}_{1,m}$$

for all k and m, $\mathfrak{J}(\mu, k, m)$ and $\tilde{\mathfrak{J}}(\mu, k, m)$ are constant for sufficiently large k. Then (3. 12) and (3. 13) imply that $\mathfrak{J}(\mu, k, m)$ and $\tilde{\mathfrak{J}}(\mu, k, m)$ are equal for sufficiently large k.

In the preceding section we needed a consequence of theorem 1:

Corollary 1. *The quantities d_m and $\chi_m(\lambda)$ remain unchanged if in $\mathbf{A}_m(\lambda)$ all blocks $A_{s-r+\tilde{r}}$ are replaced by $A_{s-r+\tilde{r}}-\alpha I$, α a complex number. The same statement holds for the defect $\hat{d}_m$ and the characterizing polynomial $\hat{\chi}_m(\lambda)$ of the matrix $\hat{\mathbf{A}}_m(\lambda)$ obtained by deleting the first ν rows and columns of $\mathbf{A}_m(\lambda)$ if the elements α_{ij}^0 of A_0 vanish for $i \leqq \nu$, $j > \nu$.*

Remark. Simple examples show that the same is not always true for the invariant polynomials of $\mathbf{A}_m(\lambda)$.

Proof. Using lemma 3. 1 and (3. 8) for two different sufficiently large k the first assertion follows immediately from the theorem. For the second part of the corollary note that all preceding considerations in this section remain valid, if the vector space $\mathfrak{O}_{k,m}$ is replaced by the vector space $\hat{\mathfrak{O}}_{k,m}$ of polynomials x in z and λ of λ-degree $< k$ and z-degree $< m$ such that the first ν components of $x|_{z=0}$ vanish.

Then one shows in an analogous manner that

$$\dim \hat{\mathfrak{N}}(\mu, k, m) = v(\mu, \hat{\chi}_m) + k\hat{d}_m = \dim \hat{\mathfrak{C}}(k, m) + \dim \hat{\mathfrak{I}}(\mu, k, m)$$

and that for sufficiently large k, the dimension of $\hat{\mathfrak{C}}(k, m)$ and the vector space $\hat{\mathfrak{I}}(\mu, k, m)$ are not changed by a substitution of $D - \alpha z^{s-r+\tilde{r}}\,\mathrm{id}_{\mathfrak{D}}$ for D. The assumption that $\alpha_{ij}^0 = 0$ if $i \leqq \nu$ and $j > \nu$ is used in the proof of the analogon of lemma 3. 1 to show the property (3. 5) of $\hat{p}_j, \hat{q}_j$ constructed from the columns of $\hat{P}(\lambda)$ and $\hat{Q}(\lambda)$.

Another important consequence of theorem 3 and lemma 3. 1 is

Corollary 2. *Let a complex number μ and a sufficiently large integer k be given. Then*

$$\dim \pi\mathfrak{N}(\mu, k, m+1) = v(\mu, \chi_{m+1}) - v(\mu, \chi_m) + k(d_{m+1} - d_m),$$

where $\pi x = x|_{z=0}$ means the projection of $\mathfrak{D}$ onto $\mathbb{C}^n[\lambda]$ with kernel $z\mathfrak{D}$.

Proof. Because of $D(zx) = z(D - z^s\,\mathrm{id}_{\mathfrak{D}})x$, $x \in \mathfrak{D}$ we obtain that

$$\dim \pi\mathfrak{N}(\mu, k, m+1) = \dim \mathfrak{N}(\mu, k, m+1) - \dim \tilde{\mathfrak{N}}(\mu, k, m),$$

where $\tilde{\mathfrak{N}}(\mu, k, m)$ is defined according to lemma 3. 1 but for $D - z^s\,\mathrm{id}_{\mathfrak{D}}$ instead of D. Because of (3. 8) and theorem 3. 1 the dimensions of $\tilde{\mathfrak{N}}(\mu, k, m)$ and $\mathfrak{N}(\mu, k, m)$ are equal and lemma 3. 1 yields

$$\dim \pi\mathfrak{N}(\mu, k, m+1) = (v(\mu, \chi_{m+1}) + kd_{m+1}) - (v(\mu, \chi_m) + kd_m).$$

Remark. Because of (3. 7) it is sufficient in corollary 2 to choose k larger than all multiplicities of μ as zero of invariant polynomials of $\mathbf{A}_{m+1}(\lambda)$ and $\mathbf{A}_m(\lambda) - \mathbf{H}_m$.

For the convenience of the reader we derive several statements of [5], sec. 3 that are used in the second section and in the sequel. Corollary 2 implies $d_{m+1} \geqq d_m$ and, because $\pi\mathfrak{N}(\mu, k, m+1)$ contains $\pi\mathfrak{N}(\mu, k, m+2)$, also yields $d_{m+2} - d_{m+1} \leqq d_{m+1} - d_m$. Since deletion of the first $n(s-r)$ rows and columns of $\mathbf{A}_m(\lambda)$ leaves a nonsingular matrix the defect satisfies $d_m \leqq n(s-r)$. Consequently there is $N \leqq n(s-r)$ such that $d_m < d_{m+1}$ for $m < N$ and $d_m = d_{m+1}$ for $m \geqq N$. For $m \geqq N$ and arbitrary k we also show $\pi\mathfrak{C}(k, m+1) = \{0\}$. If we take some $x \in \mathfrak{C}(k_0, m+1)$, then $x, \lambda x, \ldots, \lambda^{k-k_0}x$ are elements of $\mathfrak{C}(k, m+1)$ for $k > k_0$. Because $\mathfrak{N}(\mu, k, m+1)$ contains $\mathfrak{C}(k, m+1)$ and $d_{m+1} = d_m$, corollary 2 implies that $x|_{z=0} = 0$.

Note that the first part of this section is independent of the second section.

From now on we consider the case that some $\mathbf{A}_m(\lambda)$ has the maximal defect, i. e. $d_m = n(s-r)$. The above remark shows that this occurs for some $m \leqq n(s-r)$ if at all. Then the r^{th} invariant number $i^{(r)}$ for (1. 1) vanishes and because of the corollary to 2. 3 the polynomial $\chi^{(r)}(\lambda) = \dfrac{\chi_{m+1}(\lambda)}{\chi_m(\lambda)}$ has degree n. Because of corollary 2 we have $\dim \pi\mathfrak{N}(\mu, k, m+1) = v(\mu, \chi)$ for zeros μ of $\chi(\lambda)$ and sufficiently large k. So we pick some sufficiently large k (see that above remark) and for every zero μ of $\chi(\lambda)$ we choose $v(\mu, \chi)$ polynomials $u_{\mu,i}$ in $\mathfrak{D}_{k,m+1}$ such that $u_{\mu,i}|_{z=0}$, $i = 1, \ldots, v(\mu, \chi)$ are linearly independent and

$$Du_{\mu,i} \equiv 0 \bmod (\lambda - \mu)^k\mathfrak{D} + z^{m+1}\mathfrak{D} \quad (i = 1, \ldots, v(\mu, \chi)). \tag{3. 14}$$

We remark that the $u_{\mu i}$ can be chosen from the $q_{j\nu}$ constructed in the proof of lemma 3. 1. Further we write

$$u_{\mu i} = \sum_{j=1}^{k} (\lambda-\mu)^{k-j} t_{\mu ij} \quad (i=1,\ldots,v(\mu,\chi)) \tag{3. 15}$$

with polynomials $t_{\mu ij}$ independent of λ and form the matrix $T(z)$ having all $t_{\mu i1}$, μ zero of $\chi(\lambda)$ and $i=1,\ldots,v(\mu,\chi)$, as columns. Since $\chi(\lambda)$ has degree n, $T(z)$ is an n by n matrix of polynomials in z. We prove

Theorem 3. 3. *If the defect of $\mathbf{A}_m(\lambda)$ is $n(s-r)$ and $T(z)$ is constructed as above, then $T(z)$ is a meromorphic transformation of* span $\leqq m-s+r$ *and the transformation $y=T(z)\,\tilde{y}$ takes* (1. 1) *into a system of differential equations of Poincaré rank $\leqq r$.*

Here the span of $T(z)$ denotes the sum of the pole orders of $T(z)$ and $T(z)^{-1}$ at zero.

In [5], 3 it was proved that a rank reducing transformation exists if $\mathbf{A}_m(\lambda)$ has defect $n(s-r)$, but it could not be constructed from the data. Further no upper bound was known on the pole order of a rank reducing transformation. It can be shown that there is no transformation of span $< m-s+r$ reducing the Poincaré rank if m is chosen minimal, i.e. such that $d_{m-1} < d_m = d_{m+1}$.

To prove the theorem we need a rather technical lemma:

Lemma 3. 2. (i) *For all suitable μ, i, j and arbitrary nonnegative integers l there exists $u \in \mathfrak{O}$ such that*

$$D(z^l u) \equiv (\lambda-\mu)^{k+m+1} z^{s-r+l} t_{\mu ij} \bmod z^{m+1}\mathfrak{O}.$$

(ii) *If $\gamma_{\mu i}(z)$ are scalar polynomials in z, $x \in \mathfrak{O}$ and $\delta(\lambda)$ is a nontrivial polynomial in λ such that*

$$Dx \equiv \delta(\lambda) z^{s-r} \sum_{\mu,i} \gamma_{\mu i}(z)\, t_{\mu i1} \bmod z^{m+1}\mathfrak{O},$$

and if $x|_{z=0}$ vanishes, then all $\gamma_{\mu i}(0)$ vanish, too.

Proof. From (3. 14) we obtain

$$D\left((\lambda-\mu)^{j-1} u_{\mu i}\right) \equiv 0 \bmod (\lambda-\mu)^k \mathfrak{O} + z^{m+1}\mathfrak{O}$$

for $j=1,\ldots,k$ and therefore, because D increases the λ-degree at most by one,

$$D\left(\sum_{\kappa=j}^{k} (\lambda-\mu)^{k-\kappa+j-1} t_{\mu i\kappa}\right) \equiv -(\lambda-\mu)^k z^{s-r} t_{\mu ij} \bmod z^{m+1}\mathfrak{O}. \tag{3. 16}$$

Thus (i) is proved for $l=0$. By definition 3. 1 this means $z^{s-r} t_{\mu ij} \in \mathfrak{I}(\mu, k, m+1)$. (There is a little abuse of notation because the degree of $z^{s-r} t_{\mu ij}$ may exceed m.) Because of (3. 12) we obtain that $z^{s-r} t_{\mu ij} \in \mathfrak{I}(\mu, k+m+1, m+1)$, i.e. there exists $\tilde{u} \in \mathfrak{O}$ such that

$$(D - l z^s \mathrm{id}_{\mathfrak{O}})\,\tilde{u} \equiv (\lambda-\mu)^{k+m+1} z^{s-r} t_{\mu ij} \bmod z^{m+1}\mathfrak{O}.$$

Since $D(z^l \tilde{u}) \equiv z^l (D - l z^s \mathrm{id}_{\mathfrak{O}})\,\tilde{u}$, this implies (i) for $l \geqq 1$.

In the second statement we may assume that $(\lambda-\mu)^{k+m+1}$ divides $\delta(\lambda)$ for all zeros μ of $\chi(\lambda)$. Further, we only have to prove it for constant $\gamma_{\mu i}(z)$ because $\gamma_{\mu i}(z)-\gamma_{\mu i}(0)$ are linear combinations of positive z-powers and (i) yields that every summand $\delta(\lambda)z^{s-r}\frac{1}{l!}\gamma_{\mu i}^{(l)}(0)z^l t_{\mu i1}$ for $l\geqq 1$ is congruent to $D\tilde{x}$ with $\tilde{x}\in\mathfrak{D}$ satisfying $\tilde{x}|_{z=0}=0$. We may as well assume that x is a polynomial in z of degree $\leqq m$.

Now if $\gamma_{\mu i}(z)=\gamma_{\mu i}$ are constant we obtain from (3. 16) and (3. 15)

$$D\tilde{u}\equiv\sum_{\mu,i}\gamma_{\mu i}\delta(\lambda)z^{s-r}t_{\mu i1}-Dx\equiv 0 \bmod z^{m+1}\mathfrak{D},$$

where $\tilde{u}=-\sum_{\mu,i}\gamma_{\mu i}\delta(\lambda)(\lambda-\mu)^{-k}u_{\mu i}-x\in\mathfrak{D}$ is a polynomial in z of degree $\leqq m$. Thus $\tilde{u}\in\mathfrak{C}(l, m+1)$ for sufficiently large l and therefore $\pi\tilde{u}=\tilde{u}|_{z=0}$ vanishes as we remarked after corollary 2.

The assumption on x implies that

$$\sum_{\mu,i}\gamma_{\mu i}\delta(\lambda)(\lambda-\mu)^{-k}u_{\mu i}|_{z=0}=0.$$

Now for every zero μ of $\chi(\lambda)$ we can regard the left side as a polynomial in $\lambda-\mu$ with coefficients in $\mathbb{C}^n$. Since all those coefficients vanish and the summands for $\tilde{\mu}\neq\mu$ at least contain $(\lambda-\mu)^k$ we obtain using (3. 15) that $\sum_i\gamma_{\mu i}t_{\mu ij}|_{z=0}$ vanishes for all zeros μ of $\chi(\lambda)$ and $j=1,\ldots,k$. This means $\sum_i\gamma_{\mu i}u_{\mu i}|_{z=0}=0$ and, because the $u_{\mu i}|_{z=0}$ are linearly independent, all $\gamma_{\mu i}$ must vanish.

Proof of theorem 3. 2. If $T(z)$ would be singular we could find polynomials $\gamma_{\mu i}(z)$ such that not all $\gamma_{\mu i}(0)$ vanish and $\sum_{\mu i}\gamma_{\mu i}(z)t_{\mu i1}=0$. This contradicts lemma 3. 2 (ii). Now let $T(z)^{-1}=z^{-S}\hat{T}(z)$ with a power series $\hat{T}(z)$ in z, $\hat{T}(0)\neq 0$. Then $T(z)\hat{T}(z)=z^S I$. If one chooses a suitable column of $\hat{T}(z)$ and cuts sufficiently large powers of z, one finds polynomials $\gamma_{\mu i}(z)$ such that not all $\gamma_{\mu i}(0)$ vanish and

$$\sum_{\mu,i}\gamma_{\mu i}(z)t_{\mu i1}\equiv 0 \bmod z^S.$$

Because of lemma 3. 2 (ii) this is only possible for $S+s-r<m+1$, hence $\operatorname{span}T(z)\leqq S\leqq m-s+r$.

Now we show that $T(z)^{-1}t_{\mu ij}$ is a power series for all suitable μ, i, j. Choose the minimal nonnegative integer M such that $u=z^M T(z)^{-1}t_{\mu ij}$ is a power series in z and assume M is positive. Then the components of u are certain power series $\gamma_{\kappa\nu}(z)$ such that not all $\gamma_{\kappa\nu}(0)$ vanish and

$$z^M t_{\mu ij}=\sum_{\kappa,\nu}\gamma_{\kappa\nu}(z)t_{\kappa\nu1}. \tag{3. 17}$$

Because of lemma 3. 2 (i) there exists $x\in\mathfrak{D}$ such that $x|_{z=0}=0$ and

$$Dx\equiv(\lambda-\mu)^{k+m+1}z^{s-r+M}t_{\mu ij} \bmod z^{m+1}.$$

With (3. 17) this gives a contradiction to lemma 3. 2 (ii). Therefore $M=0$ and $T(z)^{-1}t_{\mu ij}$ is a power series in z.

Now (3. 14) and (3. 15) imply

$$-z^{s+1}\frac{d}{dz}t_{\mu i1}+A(z)\,t_{\mu i1}\equiv z^{s-r}(\mu t_{\mu i1}+t_{\mu i2}) \bmod z^{m+1}$$

for all zeros μ of $\chi(\lambda)$ and $i=1,\ldots,v(\mu,\chi)$. Because the pole order of $T(z)^{-1}$ does not exceed $m-s+r$ this implies

$$T(z)^{-1}\left[-z^{s+1}\frac{d}{dz}t_{\mu i1}+A(z)\,t_{\mu i1}\right]\equiv z^{s-r}\left[\mu T(z)^{-1}t_{\mu i1}+T(z)^{-1}t_{\mu i2}\right]$$

$\bmod z^{s-r+1}$. Since $T(z)^{-1}t_{\mu i\,2}$ are power series there exists an n by n matrix $\tilde{A}(z)$ of power series such that

$$T(z)^{-1}\left[-z^{s+1}\frac{d}{dz}T(z)+A(z)\,T(z)\right]=z^{s-r}\tilde{A}(z).$$

Thereby the theorem is proved.

References

[1] *W. Balser, W. B. Jurkat, D. A. Lutz*, A general theory of invariants for meromorphic differential equations. I, Formal invariants, Funk. Ekvac. **22** (1979), 197—227.
[2] *V. Dietrich*, Zur Reduktion von linearen Differentialgleichungssystemen, Math. Ann. **237** (1978), 79—95.
[3] *F. R. Gantmacher*, The theory of matrices. I, II, New York 1959.
[4] *E. L. Ince*, Ordinary differential equations, New York 1956.
[5] *D. A. Lutz, R. Schäfke*, On the identification and stability of formal invariants for singular differential equations, LA and its Appl. (to appear).
[6] *J. Moser*, The order of a singularity in Fuchs' theory, Math. Z. **72** (1960), 379—398.
[7] *F. W. Schäfke, D. Schmidt*, Gewöhnliche Differentialgleichungen, Berlin-Heidelberg 1973.
[8] *E. Wagenführer*, Über regulär-singuläre Lösungen von Systemen linearer gewöhnlicher Differentialgleichungssysteme. I, J. reine angew. Math. **267** (1974), 90—114; II, ibid. **272** (1975), 150—172.
[9] *E. Wagenführer*, On meromorphic transformations reducing the Poincaré rank of a linear differential equation at the singular point, in preparation.

Universität — Gesamthochschule Essen, Fachbereich 6 — Mathematik, Universitätsstraße 3, D-4300 Essen 1

Eingegangen 31. Januar 1985

On the Hodge-Tate decomposition in the imperfect residue field case

By *Osamu Hyodo* at Tokyo

Introduction

Let K be a complete discrete valuation field of mixed characteristics $(0, p)$ with the integer ring O_K and with the residue field $\bar{K}$, and let $\mathbb{C}_p$ be the completion of an algebraic closure of K. *Under the assumption that $\bar{K}$ is perfect*, Tate [15] calculated the continuous cohomology groups $H^q(K, \mathbb{C}_p(r))$. Namely,

$$H^q(K, \mathbb{C}_p(r)) = \begin{cases} K & \text{if } r=0 \text{ and } q=0, 1, \\ 0 & \text{otherwise.} \end{cases}$$

And he showed that the Tate module T_pG of a p-divisible group G over O_K admits a Hodge-Tate decomposition. Namely, there exists a splitting exact sequence of $\mathrm{Gal}(K_{\mathrm{sep}}/K)$-modules

$$0 \to \mathbb{C}_p(1)^s \to T_pG \otimes_{\mathbb{Z}_p} \mathbb{C}_p \to \mathbb{C}_p^t \to 0 \tag{0-1}$$

for some $s, t \geqq 0$. The existence of (0-1) is shown by using the results $H^0(K, \mathbb{C}_p) = K$ and $H^0(K, \mathbb{C}_p(r)) = 0$ for $r \neq 0$. The splitting of (0-1) follows from $H^1(K, \mathbb{C}_p(1)) = 0$.

In this paper we treat the case where $\bar{K}$ is not perfect. We shall show that the exact sequence (0-1) still exists (Remark 1 to Theorem 1), but that there are many abelian varieties defined over K whose Tate module does not admit a Hodge-Tate decomposition (i.e. (0-1) does not split) (c.f. Theorem 3). We first determine $H^q(K, \mathbb{C}_p(r))$ for all $q \geqq 0$ and $r \in \mathbb{Z}$ without the assumption that $\bar{K}$ is perfect.

Theorem 1. *Let $\hat{\Omega}^q(O_K) = \varprojlim_n \Omega_I^q(O_K)/p^n$, the p-adic completion of the q-th differential group $\Omega_I^q(O_K)$. And denote $\hat{\Omega}^q(O_K) \otimes_{O_K} K$ by $\hat{\Omega}_K^q$. Then we have for $q \geqq 0$,*

$$H^q(K, \mathbb{C}_p(r)) \cong \begin{cases} \hat{\Omega}_K^q \quad (\textit{canonically}) & \textit{if } r=q, \\ \hat{\Omega}_K^{q-1} & \textit{if } r=q-1, \\ 0 & \textit{otherwise.} \end{cases}$$

Remark 1. By Theorem 1, we have $H^0(K, \mathbb{C}_p) = K$, and $H^0(K, \mathbb{C}_p(r)) = 0$ if $r \neq 0$. So we can show that the exact sequence (0-1) still exists by the same argument as in Tate [15] § 4.

Remark 2. Let I be the index set of a p-base of $\bar{K}$. Then we have for $q \geqq 0$,

$$\hat{\Omega}^q_K \cong (\varprojlim_n (\bigwedge^q_{\mathbb{Z}} (\bigoplus_{i \in I} \mathbb{Z})) \otimes_{\mathbb{Z}} (O_K/p^n)) \otimes_{O_K} K$$

(c.f. Proposition (4-2)). In particular if $[\bar{K} : \bar{K}^p] = p^d < \infty$, then $\hat{\Omega}^q_K$ is a K-vector space of dimension $\binom{d}{q}$, so if $\bar{K}$ is perfect, $\hat{\Omega}^q_K = 0$ for $q > 0$.

Remark 3. For $q \geqq 2$, $\hat{\Omega}^q_K$ is canonically isomorphic to the completion of the q-th exterior power $(\bigwedge^q_K \hat{\Omega}^1_K)^\wedge$ by a natural topology. We shall show in § 3 that the isomorphisms of Theorem 1 are induced by the cup product, namely

$$\hat{\Omega}^q_K \cong (\bigwedge^q_K \hat{\Omega}^1_K)^\wedge \cong (\bigwedge^q_K H^1(K, \mathbb{C}_p(1)))^\wedge \xrightarrow[\cup]{\cong} H^q(K, \mathbb{C}_p(q)),$$

$$\hat{\Omega}^{q-1}_K \cong H^1(K, \mathbb{C}_p) \otimes_K H^{q-1}(K, \mathbb{C}_p(q-1)) \xrightarrow[\cup]{\cong} H^q(K, \mathbb{C}_p(q-1)),$$

where $\cup$ means the cup product.

To show that there are many abelian varieties defined over K whose p-divisible group does not admit a Hodge-Tate decomposition, we give a criterion for the Hodge-Tate decomposition in special cases (Theorem 3) by using the following Theorem 2.

Theorem 2. *Let* $\iota : \hat{K}^\times \cong H^1(K, \mathbb{Z}_p(1)) \to H^1(K, \mathbb{C}_p(1))$ *be the natural homomorphism, where* $\hat{K}^\times = \varprojlim_n K^\times/(K^\times)^{p^n}$. *(For the isomorphism* $\hat{K}^\times \cong H^1(K, \mathbb{Z}_p(1))$, *c.f. Lemma* (3-6).) *Let* $F = \bigcap_{n=1}^{\infty} \bar{K}^{p^n}$ *be the maximal perfect subfield of* $\bar{K}$, *and let* k *be the algebraic closure in* K *of the quotient field of the Witt ring* $W(F)$. *Then we have*

$$\mathrm{Ker}(\iota) = \hat{k}^\times.$$

(We shall call the above field k *the canonical subfield of* K.)

Let A be a d-dimensional abelian variety over K which has stable reduction, and let $\mathscr{A}$ be its Néron model ([10]). We assume that $\bar{K}$ is separably closed. We will consider the following two cases.

(0-2-1) The connected component of the special fiber of $\mathscr{A}$ is a power of G_m, the multiplicative group.

(0-2-2) $\mathscr{A}$ is an abelian scheme over O_K and the special fiber is an ordinary abelian variety.

Let $T_p(A)$ be the Tate module of the p-divisible group $A(p)$ attached to A. In both cases, we have the following exact sequence of $\mathrm{Gal}(K_{\mathrm{sep}}/K)$-modules:

$$0 \to \mathbb{Z}_p(1)^d \to T_p(A) \to \mathbb{Z}_p^d \to 0. \tag{0-3}$$

By Theorem 2, $T_p(A)$ admits a Hodge-Tate decomposition if and only if the image of the boundary map arising from (0-3)

$$\delta : \mathbb{Z}_p^d \to H^1(K, \mathbb{Z}_p(1)^d) \cong (\hat{K}^\times)^d$$

is contained in $(\hat{k}^\times)^d$.

To formulate our result, we consider the following condition.

(0-4) $\bar{K}$ is a separable closure of a field finitely generated over $\bar{k}$.

Theorem 3. *Under the assumption* (0-2-1) *or* (0-2-2)+(0-4), $T_p(A)$ *admits a Hodge-Tate decomposition if and only if* A *is defined over* k. *Under the assumption* (0-2-2), *if*

$$(\mathscr{A} \times_{O_K} \bar{K})(p) \ncong (G_m(p))^d \oplus (\mathbb{Q}_p/\mathbb{Z}_p)^d,$$

then $T_p(A)$ *does not admit a Hodge-Tate decomposition. Even if*

$$(\mathscr{A} \times_{O_K} \bar{K})(p) \cong (G_m(p))^d \oplus (\mathbb{Q}_p/\mathbb{Z}_p)^d,$$

$T_p(A)$ *does not necessarily admit a Hodge-Tate decomposition.*

The plan of this paper is as follows. In § 1—§ 4, we prove Theorem 1 and Theorem 2 assuming the following condition (0-5).

(0-5) K contains a primitive p^2-th root of unity, and a prime element of the canonical subfield k of K (c.f. Theorem 2) is still a prime element of K.

In § 5, we deduce the general case of Theorem 1 and Theorem 2 from the special case (0-5), by using a theorem of Epp [1]. In § 6—§ 7, we prove Theorem 3.

I am very happy to express my sincere gratitude to Professor Kazuya Kato for valuable advices. He enlightened me on the question treated in § 7. I also wish to express my hearty thanks to Professor Shuji Saito for his helpful advices and encouragements, and to my colleague Mr. Toshio Okochi for valuable discussions on this problem.

Conventions

In this paper, we fix a complete discrete valuation field K of mixed characteristics $(0, p)$. We denote by K_{sep} a separable closure of K. Choose and fix $\{\zeta_n\}_n \subset K_{\mathrm{sep}}$ such that ζ_n is a primitive p^n-th root of unity and that $\zeta_m^{p^{m-n}} = \zeta_n$ if $m \geqq n$. Let v be the additive valuation of K_{sep} such that $v(p) = 1$. For a subextension L of K_{sep}/K, we denote by O_L the integer ring of L and by $\bar{L}$ the residue field of L. We denote the r-th Tate twist by (r) (c.f. Tate [16] p. 262).

§ 1. A construction of certain subextensions of K_{sep}/K

Throughout § 1—§ 4, we always assume (0-5). In this section we construct certain subextensions of K_{sep}/K.

Choose $\{u_i\}_{i\in I}\subset O_K$ such that $\{\overline{u_i}\}_{i\in I}$ form a p-base of $\bar{K}$, and choose a p^m-th root $w_{i,m}$ of u_i for each $i\in I$ and $m\geqq 0$ satisfying $(w_{i,m+1})^p=w_{i,m}$. Let

$$K_\infty=\bigcup_{n=1}^{\infty}K(\zeta_n),\qquad L_J=\bigcup_{\substack{j\in J\\ 0<m<\infty}}K_\infty(w_{j,m})$$

for a finite subset J of I (for an empty set $\emptyset$, let $L_\emptyset=K_\infty$), and let $M=\varinjlim_J L_J$ where J runs through all finite subsets of I. Then K_∞, L_J and M are Galois extensions of K. Moreover we have isomorphisms

$$\mathrm{Gal}(L_J/K_\infty)\cong\prod_{j\in J}\mathbb{Z}_p(1)\quad\text{and}\quad\mathrm{Gal}(M/K_\infty)\cong\prod_{i\in I}\mathbb{Z}_p(1)$$

as $\mathrm{Gal}(K_\infty/K)$-modules. Choose $s_i\in\mathrm{Gal}(M/K_\infty)$ such that

$$s_i(w_{i,m})=\zeta_m w_{i,m}\quad\text{and}\quad s_i(w_{j,m})=w_{j,m}\quad\text{if } i\neq j.$$

We consider naturally s_i as an element of $\mathrm{Gal}(L_J/K_\infty)$; then $\mathrm{Gal}(L_J/K_\infty)$ is a free $\mathbb{Z}_p$-module generated by $\{s_j\}_{j\in J}$. Let σ be a generator of $\mathrm{Gal}(K_\infty/K)\cong\mathbb{Z}_p$, and choose $\tilde{\sigma}\in\mathrm{Gal}(M/K)$ such that

$$\tilde{\sigma}(\zeta_n)=\sigma(\zeta_n)\quad\text{and}\quad\tilde{\sigma}(w_{i,m})=w_{i,m}\quad\text{for all } i\in I\text{ and } m\in\mathbb{N}.$$

Then $\mathrm{Gal}(M/K)$ is generated by $\tilde{\sigma}$ and $\{s_i\}_{i\in I}$.

The advantage of the assumption that a prime element of k is still a prime element of K is that $\bar{K}_\infty/\bar{K}$ is separable, so $O_{K_\infty(w_{i,m})}$ is generated by $w_{i,m}$ over O_{K_∞}, and so we can describe explicitly the action of $\mathrm{Gal}(M/K)$ on M. To make this precise, we will give some notations. For $y\in\mathbb{Q}_p/\mathbb{Z}_p$, we define u_i^y and 1^y by

$$u_i^y=(w_{i,m})^l\quad\text{and}\quad 1^y=\zeta_m^l,$$

where $y=l\cdot p^{-m}\pmod{\mathbb{Z}_p}$ for $l\in\mathbb{Z}$, $0\leqq l<p^m$. Then we have by definition

$$\text{(1-1)}\qquad\begin{aligned}&\tilde{\sigma}(u_i^y)=u_i^y,\quad\tilde{\sigma}(1^y)=1^{\chi(\sigma)\cdot y},\\ &s_i(u_i^y)=1^y\cdot u_i^y,\quad\text{and}\quad s_i(u_j^y)=u_j^y\quad\text{if}\quad i\neq j,\end{aligned}$$

where χ is the cyclotomic character. For a finite subset $J\subset I$, let

$$B_J=\left\{\prod_{j\in J}u_j^{y_j};\ y_j\in\mathbb{Q}_p/\mathbb{Z}_p\ \text{for each } j\right\}\quad\text{and}\quad B_\emptyset=\{1\}.$$

Then B_J is a base of L_J (resp. O_{L_J}) as a K_∞-vector space (resp. free O_{K_∞}-module). For $J'\subset J$, let $A(J,J')\subset L_J$ be a sub-vector space with base $B_J\setminus B_{J'}$. Then

$$O_{A(J,J')}=\{a\in A(J,J');\ v(a)\geqq 0\}$$

is a sub-free-O_{K_∞}-module of O_{L_J} (this is not a ring). Let $D(J,J')=A(J,J')/O_{A(J,J')}$. Then by (1-1), we can show easily the following Lemma (1-2).

Lemma (1-2). *$A(J, J')$ and $O_{A(J,J')}$ are $\mathrm{Gal}(L_J/K)$-modules. Moreover for finite subsets $J'' \subset J' \subset J \subset I$, we have the following direct sum decompositions as $\mathrm{Gal}(L_J/K)$-modules*:

(1-2-1) $$A(J, J'') = A(J, J') \oplus A(J', J''),$$

(1-2-2) $$L_J = A(J, J') \oplus L_{J'},$$

(1-2-3) $$O_{A(J,J'')} = O_{A(J,J')} \oplus O_{A(J',J'')},$$

(1-2-4) $$O_{L_J} = O_{A(J,J')} \oplus O_{L_{J'}},$$

(1-2-5) $$D(J, J'') = D(J, J') \oplus D(J', J''),$$

(1-2-6) $$L_J/O_{L_J} = D(J, J') \oplus L_{J'}/O_{L_{J'}}.$$

Remark (1-3). As we have $M = \varinjlim_J L_J$ and $M/O_M = \varinjlim_J L_J/O_{L_J}$, we have

$$M/O_M = \left(\varinjlim_J D(J, \emptyset)\right) \oplus K_\infty/O_{K_\infty}$$

as $\mathrm{Gal}(M/K)$-modules.

§ 2. The determination of $H^q(M/K_\infty, M/O_M)$

In this section, we determine the groups $H^q(M/K_\infty, M/O_M)$. We denote $H^q(\mathrm{Gal}(L/K), A)$ by $H^q(L/K, A)$. We use the notations of § 1. The aim of this section is to prove the following Proposition (2-1).

Proposition (2-1). *We have*

(2-1-1) $H^q\left(M/K_\infty, \varinjlim_J D(J, \emptyset)\right)$ *is killed by p for $q \geqq 0$.*

(2-1-2) $H^q(M/K_\infty, K_\infty/O_{K_\infty}) \cong \left(\bigwedge_{\mathbb{Z}}^q \left(\bigoplus_{i \in I} \mathbb{Z}\right)\right) \otimes_{\mathbb{Z}} (K_\infty/O_{K_\infty})(-q)$ *as a $\mathrm{Gal}(K_\infty/K)$-module if $q \geqq 0$.*

The assertion (2-1-2) is immediate from the fact

$$\mathrm{Gal}(M/K_\infty) \cong \prod_{i \in I} \mathbb{Z}_p(1)$$

as a $\mathrm{Gal}(K_\infty/K)$-module. To show (2-1-1), we need the following Lemma (2-2).

Lemma (2-2). *For any finite subset $J' \subset J \subset I$ and $q \geqq 0$, $H^q\left(L_J/L_{J'}, D(J, J')\right)$ is killed by p.*

First we prove (2-1-1) by using Lemma (2-2). It suffices to show:

$$H^q\left(L_{J'}/K_\infty, H^0\left(M/L_{J'}, \varinjlim_J D(J, \emptyset)\right)\right) \text{ is killed by } p$$

for any finite subset $J' \subset I$. So it suffices to show:

$$H^q\left(L_{J'}/K_\infty, H^0\left(M/L_{J'}, D(J, \emptyset)\right)\right) \text{ is killed by } p$$

for $J' \subset J$. By (1-2) and (2-2), we have

$$H^0\left(M/L_{J'}, D(J, \emptyset)\right) = T \oplus D(J', \emptyset)$$

where $T = H^0\left(L_J/L_{J'}, D(J, J')\right)$ is killed by p. Now (2-1-1) follows from (2-2).

Next we prove (2-2) by induction on $\#(J\backslash J')$.

Step 1. We first prove (2-2) in the case $\#(J\backslash J')=1$. Let $\{j\}=J\backslash J'$. We consider naturally s_j as an element of $\mathrm{Gal}(L_J/L_{J'})$. Then s_j is a generator of $\mathrm{Gal}(L_J/L_{J'})\cong \mathbb{Z}_p$. In this case we prove a stronger result,

$$\text{(2-3)}\qquad \begin{cases} H^0(L_J/L_{J'}, D(J,J')) & \text{is killed by } p, \\ H^q(L_J/L_{J'}, D(J,J'))=0 & \text{if } q\geqq 1. \end{cases}$$

The case $q\geqq 2$ is trivial, for the cohomological dimension of $\mathbb{Z}_p$ is one. We consider the following exact sequence

$$\text{(2-4)}\qquad 0\longrightarrow H^0(L_J/L_{J'}, D(J,J'))\longrightarrow D(J,J')$$
$$\xrightarrow{1-s_j} D(J,J')\longrightarrow H^1(L_J/L_{J'}, D(J,J'))\longrightarrow 0,$$

and a commutative diagram of exact rows

$$\text{(2-5)}\qquad \begin{array}{ccccccccc} 0 & \longrightarrow & O_{A(J,J')} & \longrightarrow & A(J,J') & \longrightarrow & D(J,J') & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle 1-s_j} & & \downarrow{\scriptstyle 1-s_j} & & \downarrow{\scriptstyle 1-s_j} & & \\ 0 & \longrightarrow & O_{A(J,J')} & \longrightarrow & A(J,J') & \longrightarrow & D(J,J') & \longrightarrow & 0. \end{array}$$

Recall that $A(J,J')$ (resp. $O_{A(J,J')}$) is an $L_{J'}$-vector space (resp. a free $O_{L_{J'}}$-module) with base $\{u_j^y\}_{y\in Q_p/\mathbb{Z}_p,\, y\neq 0}$. We have by (1-1),

$$\text{(2-6)}\qquad (1-s_j)\cdot u_j^y=(1-1^y)\cdot u_j^y.$$

As $0<v(1-1^y)\leqq 1$ if $y\neq 0$, the second vertical arrow in (2-5) is an isomorphism and the cokernel of the first vertical arrow is killed by p. Now (2-3) can be seen by an easy diagram chasing.

Step 2. Assume that Lemma (2-2) is verified for $J''\subset J'\subset I$. We next prove Lemma (2-2) for $J''\subset J=J'\cup\{j\}\subset I$ for $j\notin J'$. By (1-2-5), it suffices to show that $H^q(L_J/L_{J''}, D(J,J'))$ and $H^q(L_J/L_{J''}, D(J',J''))$ are killed by p. Consider the following two spectral sequences:

$$\text{(2-7)}\qquad H^s(L_{J'}/L_{J''}, H^t(L_J/L_{J'}, D(J,J')))\Rightarrow H^{s+t}(L_J/L_{J''}, D(J,J')),$$
$$\text{(2-8)}\qquad H^s(L_{J'}/L_{J''}, H^t(L_J/L_{J'}, D(J',J'')))\Rightarrow H^{s+t}(L_J/L_{J''}, D(J',J'')).$$

Applying (2-3) to $L_J/L_{J'}$, we see by (2-7) that $H^q(L_J/L_{J''}, D(J,J'))$ is killed by p. As $\mathrm{Gal}(L_J/L_{J'})\cong\mathbb{Z}_p$, we have the following long exact sequence by (2-8):

$$\cdots\longrightarrow H^q(L_{J'}/L_{J''}, D(J',J''))\xrightarrow{\mathrm{Inf}} H^q(L_J/L_{J''}, D(J',J''))$$
$$\longrightarrow H^{q-1}(L_{J'}/L_{J''}, H^1(L_J/L_{J'}, D(J',J')))\longrightarrow H^{q+1}(L_{J'}/L_{J''}, D(J',J''))\longrightarrow\cdots.$$

The important thing is that the map "Inf" has a left inverse and consequently we have

$$\text{(2-9)}\qquad H^q(L_J/L_{J''}, D(J',J''))$$
$$\cong H^q(L_{J'}/L_{J''}, D(J',J''))\oplus H^{q-1}(L_{J'}/L_{J''}, H^1(L_J/L_{J'}, D(J',J')))$$
$$\cong H^q(L_{J'}/L_{J''}, D(J',J''))\oplus H^{q-1}(L_{J'}/L_{J''}, D(J',J'')).$$

In fact, we define

$$\rho : H^q(L_J/L_{J''}, D(J', J'')) \longrightarrow H^q(L_{J'}/L_{J''}, D(J', J''))$$

as the composite map

$$H^q(L_J/L_{J''}, D(J', J'')) \xrightarrow{\text{Res}} H^q(L_J/L_{J''\cup\{j\}}, D(J', J'')) \xleftarrow{\cong} H^q(L_{J'}/L_{J''}, D(J', J'')).$$

Here, the second isomorphism follows from the natural isomorphism

$$\mathrm{Gal}(L_J/L_{J''}) \supset \mathrm{Gal}(L_J/L_{J''\cup\{j\}}) \xrightarrow{\cong} \mathrm{Gal}(L_{J'}/L_{J''}).$$

We can see easily that $\rho \circ \mathrm{Inf}$ is the identity map. Now we see by (2-9) and the induction hypothesis that $H^q(L_J/L_{J''}, D(J', J''))$ is killed by p.

§ 3. The proof of Theorem 1 (assuming (0-5))

Keep the assumptions and the notations of § 1. The following Proposition (3-1) is shown in Tate [15] Proposition 7. (Under the assumption (0-5) we have

$$D_{K(\zeta_{n+1})/K(\zeta_n)} = O_K(\zeta_{n+1}) \otimes_{O_k(\zeta_{n+1})} D_{k(\zeta_{n+1})/k(\zeta_n)},$$

where D denotes the relative different. So the same argument as in [15] § (3. 1) works.)

Proposition (3-1). (*The determination of* $H^q(K_\infty/K, K_\infty/O_{K_\infty}(r))$.)

(3-1-1) *The kernel and the cokernel of the natural homomorphisms*

$$H^q(K_\infty/K, K/O_K) \longrightarrow H^q(K_\infty/K, K_\infty/O_{K_\infty}) \quad (q=0, 1)$$

are killed by some power of p.

(3-1-2) $H^q(K_\infty/K, K_\infty/O_{K_\infty}(r))$ *is killed by some power of* p *for* $q=0, 1$ *if* $r \neq 0$.

(3-1-3) $H^q(K_\infty/K, K_\infty/O_{K_\infty}(r))=0$ *for all* $r \in \mathbb{Z}$ *if* $q \geqq 2$.

By (1-3), (2-1), (3-1), and the spectral sequence

$$H^s(K_\infty/K, H^t(M/K_\infty, M/p^n O_M)(r)) \Rightarrow H^{s+t}(M/K, M/p^n O_M(r)),$$

we have the following Proposition (3-2).

Proposition (3-2). (*The determination of* $H^q(M/K, M/p^n O_M(r))$.)

(3-2-1) *The kernel and the cokernel of the natural homomorphisms*

$$H^q(M/K, M/p^n O_M(q)) \longrightarrow H^0\left(K_\infty/K, \bigwedge_{\mathbb{Z}}^q \left(\bigoplus_{i\in I} \mathbb{Z}\right) \otimes_{\mathbb{Z}} (K_\infty/p^n O_{K_\infty})\right)$$

$$H^0\left(K_\infty/K, \bigwedge_{\mathbb{Z}}^q \left(\bigoplus_{i\in I} \mathbb{Z}\right) \otimes_{\mathbb{Z}} (K/p^n O_K)\right) \longrightarrow H^0\left(K_\infty/K, \bigwedge_{\mathbb{Z}}^q \left(\bigoplus_{i\in I} \mathbb{Z}\right) \otimes_{\mathbb{Z}} (K_\infty/p^n O_{K_\infty})\right)$$

$$H^1\left(K_\infty/K, \bigwedge_{\mathbb{Z}}^q \left(\bigoplus_{i\in I} \mathbb{Z}\right) \otimes_{\mathbb{Z}} (K/p^n O_K)\right) \longrightarrow H^{q+1}(K, M/p^n O_M(q))$$

are killed by some power of p *which is independent of* n.

(3-2-2) *In the case* $r \neq q$, $q-1$ *or* $(q, r)=(0, -1)$, $H^q(M/K, M/p^n O_M(r))$ *is killed by some power of* p *which is independent of* n.

Proposition (3-3). *Let* $\hat{M}$ *be the completion of* M, *and let*

$$O_K\hat{}(I, q)=\varprojlim_n\left(\bigwedge_{\mathbb{Z}}^q\left(\bigoplus_{i\in I}\mathbb{Z}\right)\right)\otimes_{\mathbb{Z}}(O_K/p^n).$$

Then we have

$$H^0(M/K, \hat{M}(r))\cong\begin{cases}K & \text{if } r=0,\\ 0 & \text{otherwise.}\end{cases}$$

And we have for $q\geqq 1$,

$$H^q(M/K, \hat{M}(r))\cong\begin{cases}O_K\hat{}(I, q)\otimes_{O_K}K & \text{if } r=q,\\ O_K\hat{}(I, q-1)\otimes_{O_K}K & \text{if } r=q-1,\\ 0 & \text{otherwise.}\end{cases}$$

For $q\geqq 2$, *the isomorphisms above are induced by the cup product, namely we have*

$$O_K\hat{}(I, q)\otimes_{O_K}K\cong\left(\bigwedge_K^q H^1(M/K, \hat{M}(1))\right)\hat{}\xrightarrow[\upsilon]{\cong}H^q(M/K, \hat{M}(q)),$$

$$O_K\hat{}(I, q-1)\otimes_{O_K}K\cong H^1(M/K, \hat{M})\otimes_K H^{q-1}(M/K, \hat{M}(q-1))\xrightarrow[\upsilon]{\cong}H^q(M/K, \hat{M}(q-1)),$$

where υ *means the cup product.*

By Proposition (3-2) and Lemma (3-6) below, we have for $q\geqq 1$,

$$H^q(M/K, \hat{M}(r))\cong H^q(M/K, K(r))$$

$$\cong\begin{cases}H^0(K_\infty/K, H^q(M/K_\infty, K)(q))\cong O_K\hat{}(I, q)\otimes_{O_K}K & \text{if } r=q,\\ H^1(K_\infty/K, H^{q-1}(M/K_\infty, K)(q-1))\cong O_K\hat{}(I, q-1)\otimes_{O_K}K & \text{if } r=q-1,\\ 0 & \text{otherwise.}\end{cases}$$

Here $H^q(M/K, K(r))$ is defined for the topology of K which is induced from its valuation. Now Proposition (3-3) follows from the fact that the cup product induces isomorphisms

$$\left(\bigwedge_K^q H^1(M/K, K(1))\right)\hat{}\xrightarrow[\upsilon]{\cong}H^q(M/K, K(q))$$

$$H^1(M/K, K)\otimes_K H^{q-1}(M/K, K(q-1))\xrightarrow[\upsilon]{\cong}H^q(M/K, K(q-1)).$$

which can be seen from the structure of $\mathrm{Gal}(M/K)$.

Remark (3-4). Here we give an explicit description of the isomorphism

$$H^1(M/K, \hat{M}(1))\cong O_K\hat{}(I, 1)\otimes_{O_K}K.$$

Let $\{b_i\}_{i\in I}$ be the canonical base of $O_K\hat{}(I, 1)$, and let f_i be the 1-cocycle of $\mathrm{Gal}(M/K)$ with coefficients in $\hat{M}(1)$ characterised by

$$f_i(s_i)=1,\quad f_i(\tilde{\sigma})=0,\quad \text{and}\quad f_i(s_j)=0\quad \text{if}\quad i\neq j.$$

Then the above isomorphism is given by

$$(\text{the class of } f_i)\longmapsto b_i\otimes 1.$$

Now Theorem 1 under the assumption (0-5) is immediate from Proposition (3-3), Lemma (3-5) below and the fact that cup products commute with inflation maps.

Lemma (3-5). *The inflation map*

$$H^q(M/K, \hat{M}(r)) \longrightarrow H^q(K, \mathbb{C}_p(r))$$

is an isomorphism for any $q \geqq 0$.

Lemma (3-5) is a direct consequence of the following results (3-5-1) and (3-5-2) of Tate [15] Corollary 2 to Proposition 9 which states that $H^q(M, O_{K_{\text{sep}}})$ is killed by p for $q \geqq 1$ with the *discrete* topology in $O_{K_{\text{sep}}}$. (As M is a $\mathbb{Z}_p$-extension of a henselian discrete valuation field $K(w_{i,m};\ i \in I,\ m \in \mathbb{N})$ with perfect residue field, the same argument as in [15] § (3. 2) works.)

(3-5-1) The cokernel of the natural injection $M/O_M \to H^0(M, K_{\text{sep}}/O_{K_{\text{sep}}})$ is killed by p.

(3-5-2) $H^q(M, K_{\text{sep}}/O_{K_{\text{sep}}})$ is killed by p for $q \geqq 1$.

Lemma (3-6) (Tate [16] Proposition (2. 2)). *Let A be a G-module with sub-G-module B satisfying $\varprojlim_n A/p^n B \cong A$. Then we have*

$$H^0(G, A) = \varprojlim_n H^0(G, A/p^n B).$$

And we have the exact sequence

$$0 \longrightarrow \varprojlim_n{}^1 H^{q-1}(G, A/p^n B) \longrightarrow H^q(G, A) \longrightarrow \varprojlim_n H^q(G, A/p^n B) \longrightarrow 0.$$

In particular if the system $\{H^{q-1}(G, A/p^n B), f_{m,n}\}$ ($f_{m,n}$ is induced by the natural map $A/p^m B \longrightarrow A/p^n B$ for $m > n$) satisfies the condition of Mittag-Leffler, then we have

$$H^q(G, A) \cong \varprojlim_n H^q(G, A/p^n B).$$

Lemma (3-6) is slightly different from Tate [16] Proposition (2. 2). But it can be seen easily that the same proof works.

Remark (3-7). We shall show in § 4 that $\hat{O_K}(I, q) \otimes_{O_K} K$ is isomorphic to $\hat{\Omega}^q_K$ and there exists a canonical isomorphism $H^q(K, \mathbb{C}_p(q)) \cong \hat{\Omega}^q_K$.

Remark (3-8). We can see by Proposition (3-2) and the proof of (3-5) that the system $\{H^q(K, K_{\text{sep}}/p^n O_{K_{\text{sep}}}(r))\}_n$ satisfies the condition of Mittag-Leffler. This result will be used in the proof of Lemma (5-1).

§ 4. The canonical isomorphism and the proof of Theorem 2 (assuming (0-5))

In this section we will prove the existence of a canonical isomorphism

$$H^q(K, \mathbb{C}_p(q)) \cong \hat{\Omega}^q_K,$$

and give the proof of Theorem 2 assuming (0-5). Note that it suffices to show the canonical isomorphism for $q = 1$.

We give some notations. For a ring B and a ring C over B, let $\Omega_B(C)$ be the module of relative differential forms and let $\hat{\Omega}_B(C)=\varprojlim_n \Omega_B(C)/p^n$ be the p-adic completion of $\Omega_B(C)$. We write simply $\Omega(C)$ for $\Omega_{\mathbf{Z}}(C)$ and $\hat{\Omega}(C)$ for $\hat{\Omega}_{\mathbf{Z}}(C)$. Let $\bar{\Omega}_K=\varinjlim \hat{\Omega}(O_L)$ where L runs through all finite subextensions of K_{sep}/K. In this section we write $\hat{\Omega}(O_K)$ (resp. $O_K^{\wedge}(I)$) instead of $\hat{\Omega}^1(O_K)$ (resp. $O_K^{\wedge}(I,1)$) for simplicity.

Let K_0 be a subfield of K such that

(4-1-1) K_0 is complete with the discrete valuation induced from K.

(4-1-2) The residue field $\overline{K_0}$ coincides with $\bar{K}$.

(4-1-3) p is a prime element of K_0.

(Such a subfield K_0 exists (not necessarily unique) by Theorem (31. 1) of Nagata [10].) Then we can choose $\{u_i\}_{i\in I}$ in § 1 as a subset of O_{K_0}, so in the following we assume $\{u_i\}_{i\in I}\subset O_{K_0}$. The following Proposition (4-2) in the case where $\bar{K}$ is perfect was proved in Fontaine [2] Theorem 1.

Proposition (4-2). (4-2-1) *Let b_i be as in Remark* (3-4). *There is an exact sequence of O_K-modules*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & O_K^{\wedge}(I) & \longrightarrow & \hat{\Omega}(O_K) & \longrightarrow & \Omega_{O_{K_0}}(O_K) & \longrightarrow & 0. \\ & & \cup & & \cup & & & & \\ & & b_i & \longmapsto & du_i/u_i & & & & \end{array}$$

In particular, we have $O_K^{\wedge}(I)\otimes_{O_K} K\cong \hat{\Omega}(O_K)\otimes_{O_K} K$.

(4-2-2) *There is an exact sequence of* $\mathrm{Gal}(K_{\mathrm{sep}}/K)$-modules

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K_{\mathrm{sep}}/\mathfrak{d}(1) & \longrightarrow & \bar{\Omega}_K & \longrightarrow & O_K^{\wedge}(I)\otimes_{O_K} K_{\mathrm{sep}} & \longrightarrow & 0, \\ & & \cup & & \cup & & & & \\ & & x\otimes(\zeta_n)_n & \longmapsto & (p^n x\cdot d\zeta_n)/\zeta_n & & & & \end{array}$$

where $\mathfrak{d}=\left\{x\in K_{\mathrm{sep}} \mid v(x)\geqq -\dfrac{1}{p-1}\right\}$. *This sequence splits as a sequence of $O_{K_{\mathrm{sep}}}$-modules. In fact, the splitting map $O_K^{\wedge}(I)\otimes_{O_K} K_{\mathrm{sep}}\longrightarrow \bar{\Omega}_K$ is given by*

$$b_i\otimes 1\longmapsto du_i/u_i=p^j\cdot dw_{i,j}/w_{i,j}.$$

We leave the proof of Proposition (4-2) to the end of this section. We go forward to the proof of the existence of a canonical isomorphism. The following construction was suggested by K. Kato. Consider the following diagram of exact sequences.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & {}_{p^{n+1}}\bar{\Omega}_K & \longrightarrow & \bar{\Omega}_K & \xrightarrow{p^{n+1}} & \bar{\Omega}_K & \longrightarrow & 0 \\ & & \downarrow p & & \downarrow p & & \downarrow \mathrm{id.} & & \\ 0 & \longrightarrow & {}_{p^n}\bar{\Omega}_K & \longrightarrow & \bar{\Omega}_K & \xrightarrow{p^n} & \bar{\Omega}_K & \longrightarrow & 0. \end{array}$$

Take the inverse limit and tensor $\otimes_{O_K} K$. Then by (4-2-2), we have the exact sequence

$$0 \longrightarrow \mathbb{C}_p(1) \longrightarrow D \otimes_{O_K} K \longrightarrow \bar{\Omega}_K \otimes_{O_K} K \longrightarrow 0, \tag{4-3}$$

where $D = \varprojlim (\xrightarrow{p} \bar{\Omega}_K \xrightarrow{p} \bar{\Omega}_K \xrightarrow{p} \bar{\Omega}_K)$. *Now we define the canonical homomorphism*

$$\hat{\Omega}^1_K \cong \hat{\Omega}(O_K) \otimes_{O_K} K \longrightarrow H^1(K, \mathbb{C}_p(1))$$

to be the composite map $\delta \circ h$ *where* h *is the inclusion map*

$$\hat{\Omega}(O_K) \otimes_{O_K} K \longrightarrow H^0(K, \bar{\Omega}_K \otimes_{O_K} K)$$

and δ *is the connecting homomorphism arising from the exact sequence* (4-3). We can check easily that the composition

$$\hat{O_K}(I) \otimes_{O_K} K \overset{\alpha}{\cong} \hat{\Omega}(O_K) \otimes_{O_K} K \xrightarrow{\delta \circ h} H^1(K, \mathbb{C}_p(1)) \overset{\beta}{\cong} \hat{O_K}(I) \otimes_{O_K} K$$

where α and β are defined in (4-2-1), (3-4) and (3-5) respectively, is the identity map. Thus $\delta \circ h$ is an isomorphism.

Next we prove Theorem 2 by using the canonical isomorphism. Theorem 2 is a direct consequence of the commutativity of the following diagram:

$$\begin{array}{ccc} H^1(K, \mathbb{Z}_p(1)) & \xrightarrow{\iota} & H^1(K, \mathbb{C}_p(1)) \\ \uparrow & & \uparrow \\ \hat{K}^\times & \cong & \text{canonical isomorphism} \\ \downarrow \text{dlog} & & \\ \hat{\Omega}_{O_k}(O_K) \otimes_{O_K} K \cong \hat{\Omega}(O_K) \otimes_{O_K} K. & & \end{array}$$

Here, the homomorphism dlog is induced by

$$K^\times \to \Omega_{O_k}(O_K); \quad \begin{cases} \pi_k \mapsto 0, \\ x \mapsto dx/x \quad \text{for} \quad x \in O_K^\times, \end{cases}$$

and the isomorphism $\hat{\Omega}_{O_k}(O_K) \otimes_{O_K} K \cong \hat{\Omega}(O_K) \otimes_{O_K} K$ follows from the exact sequence

$$O_K \otimes_{O_k} \Omega(O_k) \longrightarrow \Omega(O_K) \longrightarrow \Omega_{O_k}(O_K) \longrightarrow 0$$

and the fact that $\hat{\Omega}(O_k) \cong \Omega_{O_{k_0}}(O_k)$ is killed by $\mathfrak{D}_{k/k_0}$, the different of k/k_0 (Proposition (4-2-1) and Serre [13] III Proposition 14). Here $k_0 = K_0 \cap k$. And the commutativity of the diagram can be seen from the following commutative diagram of exact sequences, where the vertical arrows are defined by $x \mapsto dx/x$.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z}/p^n\mathbb{Z}(1) & \longrightarrow & O_{K_{\text{sep}}}^\times & \xrightarrow{p^n} & O_{K_{\text{sep}}}^\times & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & {}_{p^n}\bar{\Omega}_K & \longrightarrow & \bar{\Omega}_K & \xrightarrow{p^n} & \bar{\Omega}_K & \longrightarrow & 0. \end{array}$$

Now we return to the proof of Proposition (4-2). The key fact is the following Lemma (4-4).

Lemma (4-4). *We have* $\hat{\Omega}(O_{K_0}) \cong \hat{O_{K_0}}(I)$.

As p is a prime element of K_0, we have $\Omega(O_{K_0})/p \cong \Omega(\overline{K_0})$. Note that $\Omega(\overline{K_0}) = \Omega(\bar{K})$ is a vector space over $\bar{K}$ with base $\{\overline{du_i}\}_{i \in I}$. We have inductively

$$\Omega(O_{K_0})/p^n \cong \bigoplus_{i \in I} (O_{K_0}/p^n) \cdot du_i \tag{4-5}$$

by the fact that $du_i \in \Omega(O_{K_0})$ is torsion free and by using the exact sequence

$$\Omega(\bar{K}) \longrightarrow \Omega(O_{K_0})/p^{n+1} \longrightarrow \Omega(O_{K_0})/p^n \longrightarrow 0.$$

Now Lemma (4-4) follows from (4-5).

First we prove (4-2-1). Consider the following exact sequence:

$$(O_K \otimes_{O_{K_0}} \Omega(O_{K_0}))/p^n \longrightarrow \Omega(O_K)/p^n \longrightarrow \Omega_{O_{K_0}}(O_K) \longrightarrow 0, \tag{4-6}$$

where n is a sufficiently large integer such that $p^n \Omega_{O_{K_0}}(O_K) = 0$. The kernel of the first arrow of this sequence is killed by $\mathfrak{D}_{K/K_0}$ because du_i is torsion free in $\Omega(O_K)$. Now (4-2-1) follows from Lemma (4-4) and (4-6). Next we prove (4-2-2). Let $M_0 = K_0(w_{i,m};\ i \in I,\ m \in \mathbb{N})$. We can see easily that

$$\begin{array}{ccc} \varinjlim \hat{\Omega}(L) & \cong & O_{K_0}(I) \otimes_{O_{K_0}} M_0, \\ \cup & & \cup \\ b_i & \mapsto & du_i/u_i \end{array}$$

where L runs through all finite subextensions of M_0/K_0.

Lemma (4-7). *Let K' be a finite extension of K_0, $M' = K'M_0$, and let*

$$\bar{\Omega}_{K'}(M') = \varinjlim \hat{\Omega}(L),$$

where L runs through all finite subextensions of M'/K'. Then we have the following exact sequence of $O_{M'}$-modules:

$$0 \longrightarrow O_{M'} \otimes_{O_{M_0}} \bar{\Omega}_{K_0}(M_0) \longrightarrow \bar{\Omega}_{K'}(M') \longrightarrow \Omega_{O_{M_0}}(O_{M'}) \longrightarrow 0.$$

For the proof of Lemma (4-7), it suffices to consider the following two cases.

(4-8-1) The case where M'/M_0 is unramified. For a sufficiently large subextension L of M_0/K_0, $K'L/L$ becomes an unramified extension. Fix such L and let $L' = K'L$. As we can choose the p-base of $\overline{L'}$ to be the same as the one of $\bar{L}$, we have by Lemma (4-4), $\hat{\Omega}(O_{L'}) \cong O_{L'} \otimes_{O_L} \hat{\Omega}(O_L)$. Now Lemma (4-7) follows from the fact that $\Omega_{O_{M_0}}(O_{M'}) = 0$.

(4-8-2) The case where M'/M_0 is totally ramified. Note that $\overline{M_0}$ is perfect. For a sufficiently large subextension L of M_0/K_0, $K'L/L$ becomes a totally ramified extension. Let $L' = K'L$. By the same argument as in the proof of (4-2-1), we have the following exact sequence:

$$0 \longrightarrow O_{L'} \otimes_{O_L} \hat{\Omega}(O_L) \longrightarrow \hat{\Omega}(O_{L'}) \longrightarrow \Omega_{O_L}(O_{L'}) \longrightarrow 0. \tag{4-8-2-1}$$

As $\pi_{L'}$ generates $O_{M'}$ over O_{M_0}, we have

$$\Omega_{O_{M_0}}(O_{M'}) \cong O_{M'} \otimes_{O_{L'}} \Omega_{O_L}(O_{L'}). \tag{4-8-2-2}$$

Lemma (4-7) follows from (4-8-1) and (4-8-2).

By Lemma (4-7), we have the following exact sequence of $O_{K_{\mathrm{sep}}}$-modules:

$$(4\text{-}9)\qquad 0 \longrightarrow O_{K_{\mathrm{sep}}} \otimes_{O_{M_0}} \bar{\Omega}_{K_0}(M_0) \longrightarrow \bar{\Omega}_{K_0}(K_{\mathrm{sep}}) \longrightarrow \Omega_{O_{M_0}}(O_{K_{\mathrm{sep}}}) \longrightarrow 0.$$

$$\bar{\Omega}_{K_0}(K_{\mathrm{sep}}) = \bar{\Omega}_K$$

By Fontaine [2] Theorem 1, we have

$$\begin{array}{ccc} K_{\mathrm{sep}}/\vartheta(1) & \cong & \Omega_{O_{M_0}}(O_{K_{\mathrm{sep}}}) \\ \cup & & \cup \\ x \otimes (\zeta_n)_n & \mapsto & p^n x \cdot d\zeta_n/\zeta_n \end{array}$$

where $\vartheta = \left\{ x \in K_{\mathrm{sep}} \mid v(x) \geqq -\dfrac{1}{p-1} \right\}$. Using the above isomorphism, we see that the exact sequence (4-9) splits as a sequence of $O_{K_{\mathrm{sep}}}$-modules. Now (4-2-2) can be seen from the following explicit description of the action of $\mathrm{Gal}(K_{\mathrm{sep}}/K)$.

(4-10-1) For $\tau \in \mathrm{Gal}(K_{\mathrm{sep}}/K)$, $\tau(d\zeta_n/\zeta_n) = \chi(\tau) \cdot d\zeta_n/\zeta_n$, where χ is the cyclotomic character, i.e. $\tau(\zeta_n) = \zeta_n^{\chi(\tau)}$.

(4-10-2) $\tilde{s}_i(dw_{i,m}/w_{i,m}) = d\zeta_m/\zeta_m + dw_{i,m}/w_{i,m}$, where the image of $\tilde{s}_i$ in $\mathrm{Gal}(M/K)$ coincides with s_i.

Remark (4-11). Proposition (4-2) is still valid without the assumption (0-5), as is seen from its proof.

§ 5. The proof of Theorem 1 and Theorem 2 (for the general case)

In this section we prove Theorem 1 and Theorem 2 in general. By a theorem of Epp [1], there is a finite Galois extension K' of K satisfying the condition (0-5). Note that $\hat{\Omega}^q_{K'} \cong K' \otimes_K \hat{\Omega}^q_K$ ((4-2-1) and Remark (4-11)). Theorem 1 follows from Lemma (5-1) below and the fact that cup products commute with restriction maps.

Lemma (5-1). *Let the notations be as above. Then the restriction map*

$$H^q(K, \mathbb{C}_p(r)) \longrightarrow H^0(K'/K, H^q(K', \mathbb{C}_p(r)))$$

is an isomorphism.

The surjectivity can be seen from Lemma (3-6) and the following fact:

$$\begin{aligned} \varprojlim_n H^q(K, K_{\mathrm{sep}}/p^n O_{K_{\mathrm{sep}}}(r)) &\cong \varprojlim_n H^0(K'/K, H^q(K', K_{\mathrm{sep}}/p^n O_{K_{\mathrm{sep}}}(r))) \\ &\cong H^0(K'/K, H^q(K', \mathbb{C}_p(r))) \end{aligned}$$

which is a consequence of Lemma (3-6), Remark (3-8) and the spectral sequence

$$H^s(K'/K, H^t(K', K_{\mathrm{sep}}/p^n O_{K_{\mathrm{sep}}}(r))) \Rightarrow H^{s+t}(K, K_{\mathrm{sep}}/p^n O_{K_{\mathrm{sep}}}(r)).$$

The injectivity follows from the well-known formula

$$\mathrm{Cor} \circ \mathrm{Res} = [K' : K] \cdot \mathrm{id}.$$

We next prove Theorem 2. Let k' be the canonical subfield of K'. Note that we may assume $\mathrm{Gal}(K'/K) \cong \mathrm{Gal}(k'/k)$. (This is also the content of a theorem of Epp.) As it is easy to see $H^0(K'/K, (K')^\times) = \hat{K}^\times$ and $H^0(k'/k, (k')^\times) = \hat{k}^\times$, Theorem 2 can be seen by taking the $\mathrm{Gal}(K'/K)$-invariant of the following exact sequence (5-2):

$$0 \longrightarrow (k')^\times \longrightarrow (K')^\times \longrightarrow H^1(K', \mathbb{C}_p(1)). \tag{5-2}$$

§ 6. Ordinary abelian varieties (a preparation for § 7)

This section is a preparation for § 7. Let F be an algebraically closed field such that $\mathrm{ch}(F) = p$, and let E be a finitely generated extension of F. We fix an ordinary abelian variety A of dimension d over E. Let $A(p)$ denote the p-divisible group attached to A and let $G_m(p)$ denote the p-divisible group attached to the multiplicative group G_m. In this section we will prove the following Proposition (6-1).

Proposition (6-1). *Assume that* $A(p) \cong (G_m(p))^d \oplus (\mathbb{Q}_p/\mathbb{Z}_p)^d$ *as a p-divisible group over* E_{sep}. *Then* $A \times_E E_{\mathrm{sep}}$ *is defined over* F.

First we recall a result of Serre-Tate [5].

Proposition (6-2). *Let* R *be an Artinian local ring with residue field* $\bar{R}$ *of characteristic* p. *There is an equivalence of categories* $C_1 \xrightarrow{\cong} C_2$, *where*:

C_1 *is the category of abelian schemes over* R.

C_2 *is the category of pairs* (B, X), *where* B *is an abelian scheme over* $\bar{R}$, *and where* X *is a p-divisible group over* R *whose reduction* $\bar{X}$ *coincides with the p-divisible group* $B(p)$ *attached to* B. (*The functor* $C_1 \to C_2$ *is the obvious one.*)

Lemma (6-3). *Let* $C = \mathrm{Spec}\, D$ *be an affine scheme such that* $\mathrm{Pic}\, C = 0$, *and let* $\hat{D}^\times = \varprojlim_n D^\times/p^n$ *be the p-adic completion of* $D^\times$. *Then we have*

$$\mathrm{Ext}^1_C((\mathbb{Q}_p/\mathbb{Z}_p)^m, (G_m(p))^n) \cong \varprojlim_r H^1_{fl}(C, (\mathbb{Z}/p^r\mathbb{Z}(1))^{mn}) \cong (\hat{D}^\times)^{mn}.$$

Proof. At first, we show that $\mathscr{E}xt^1_C((\mathbb{Z}/p^r\mathbb{Z})^m, (G_m(p))^n) = 0$. In fact, from the exact sequence

$$0 \longrightarrow \mathbb{Z}^m \xrightarrow{p^r} \mathbb{Z}^m \longrightarrow (\mathbb{Z}/p^r\mathbb{Z})^m \longrightarrow 0,$$

we have

$$(G_m(p))^{mn} \xrightarrow{p^r} (G_m(p))^{mn} \longrightarrow \mathscr{E}xt^1_C((\mathbb{Z}/p^r\mathbb{Z})^m, (G_m(p))^n) \longrightarrow 0.$$

Now the first isomorphism follows from the spectral sequence

$$H^p_{fl}(C, \mathscr{E}xt^q_C((\mathbb{Z}/p^r\mathbb{Z})^m, (G_m(p))^n)) \Rightarrow \mathrm{Ext}^{p+q}_C((\mathbb{Z}/p^r\mathbb{Z})^m, (G_m(p))^n),$$

and

$$\mathrm{Ext}^1_C((\mathbb{Q}_p/\mathbb{Z}_p)^m, (G_m(p))^n) \cong \varprojlim_r \mathrm{Ext}^1_C((\mathbb{Z}/p^r\mathbb{Z})^m, (G_m(p))^n).$$

The second isomorphism follows from the exact sequence of flat sheaves

$$0 \longrightarrow \mathbb{Z}/p^r\mathbb{Z}(1) \longrightarrow G_m \xrightarrow{p^r} G_m \longrightarrow 0,$$

and the fact that $H^1_{fl}(C, G_m) = \mathrm{Pic}\, C = 0$. (Q.E.D.)

We will describe the isomorphism we have just proved,

$$Q : \mathrm{Ext}^1_C((\mathbb{Q}_p/\mathbb{Z}_p)^m, (G_m(p))^n) \xrightarrow{\cong} (\hat{D}^\times)^{mn}.$$

Let $Y \in \mathrm{Ext}^1_C((\mathbb{Q}_p/\mathbb{Z}_p)^m, (G_m(p))^n)$ and let Y_r be the subgroup scheme of Y killed by p^r. Then we have

$$0 \longrightarrow (\mathbb{Z}/p^r\mathbb{Z}(1))^n \longrightarrow Y_r \longrightarrow (\mathbb{Z}/p^r\mathbb{Z})^m \longrightarrow 0. \tag{6-4}$$

Let $(q_r^{i,1}, \ldots, q_r^{i,n})$ be the image of $(\ldots, 0, \overset{(i)}{1}, 0, \ldots)$ by the boundary map arising from (6-4),

$$\delta : (\mathbb{Z}/p^r\mathbb{Z})^m \longrightarrow H^1_{fl}(C, (\mathbb{Z}/p^r\mathbb{Z}(1))^n) \cong (D^\times/(D^\times)^{p^r})^n.$$

Let $q_Y^{i,j} \in \hat{D}^\times$ such that $q_r^{i,j} \equiv q_Y^{i,j} \pmod{(D^\times)^{p^r}}$ for all r. Then

$$Q(Y) = \{q_Y^{i,j}\}_{1 \leqq i \leqq m,\, 1 \leqq j \leqq n}.$$

It can be easily seen that $Y \cong (G_m(p))^n \oplus (\mathbb{Q}_p/\mathbb{Z}_p)^m$ if and only if $q_Y^{i,j} = 1$ for all i, j.

Now we prove Proposition (6-1). We fix an ample divisor $\mathscr{D}$ of A. We replace E by a sufficiently large finite subextension of E_{sep}/E such that

$$_3A(E) \cong (\mathbb{Z}/3\mathbb{Z})^{2d} \quad \text{in the case} \quad p \neq 3, \tag{6-5-1}$$

$$_4A(E) \cong (\mathbb{Z}/4\mathbb{Z})^{2d} \quad \text{in the case} \quad p = 3. \tag{6-5-2}$$

There are an abelian scheme $\tilde{A} \to S$ and an ample divisor $\tilde{\mathscr{D}}$ of $\tilde{A}$ satisfying the following properties:

(6-6-1) S is a smooth integral scheme over F and the function field of S is isomorphic to E.

$$A \cong \tilde{A} \times_S E \text{ and } \mathscr{D} \cong \tilde{\mathscr{D}} \times_S E. \tag{6-6-2}$$

(6-6-3) $\tilde{A} \times_S \bar{s}$ is an ordinary abelian variety for any geometric point $\bar{s} \to S$.

$$_3\tilde{A}(S) \cong (\mathbb{Z}/3\mathbb{Z})^{2d} \quad \text{in the case} \quad p \neq 3. \tag{6-6-4-1}$$

$$_4\tilde{A}(S) \cong (\mathbb{Z}/4\mathbb{Z})^{2d} \quad \text{in the case} \quad p = 3. \tag{6-6-4-2}$$

Fix a closed point s of S. Note that the residue field of $O_{S,s}$ is F. Let $\mathscr{A} = \tilde{A} \times_S \hat{O}_{S,s}$ where $\hat{O}_{S,s}$ is the completion of $O_{S,s}$ with respect to the maximal ideal. By Lemma (6-3) and by the fact that

$$\varprojlim_n O^\times_{S,s}/p^n \longrightarrow \varprojlim_n E^\times_{\mathrm{sep}}/p^n$$

is injective, we have

$$\mathscr{A}(p) \cong (\mathbb{Q}_p/\mathbb{Z}_p)^d \oplus (G_m(p))^d \quad \text{over} \quad \hat{O}_{S,s}.$$

Let $B = \tilde{A} \times_S s$ be the abelian variety over F. In Proposition (6-2), we can replace "Artinian local ring" by "complete noetherian local ring" and "scheme" by "formal scheme". So by Proposition (6-2), we have $\mathscr{A} \cong B \times_F \hat{O}_{S,s}$. As we have $\mathrm{Hom}(\mathscr{A}, \mathscr{A}) = \mathrm{Hom}(B, B)$ by Mumford [9] Corollary 6.2., we see that the polarization

$$\lambda : \mathscr{A} \to \mathscr{A}$$

associated to $\tilde{\mathscr{D}} \times_S \hat{O}_{S,s}$ is also defined over F (by $\tilde{\mathscr{D}} \times_S s$). Now we see that A is defined over F by the existence of the fine moduli scheme with level n structure for $n \geqq 3$ (Mumford [9] p. 139).

§ 7. Some criterion for the Hodge-Tate decomposition

In this section we prove Theorem 3.

First we consider the case (0-2-1). We use the following result.

Proposition (7-1) (Raynaud [12]). *There is a discrete subgroup $\Lambda \subset (K^\times)^d$ isomorphic to $\mathbb{Z}^d$, such that $A \cong (K^\times)^d/\Lambda$ as a rigid analytic group.*

Let $\{\alpha_i = (a_{i,1}, \ldots, a_{i,d}) \in (K^\times)^d\}_{1 \leq i \leq d}$ be a base of Λ. A straightforward calculation shows that

$$\delta(\ldots, 0, \overset{(i)}{1}, 0, \ldots) = \alpha_i$$

(δ is defined in the Introduction) for a suitable choice of a base of $T_p(A)$. Now the assertion that $\delta(\mathbb{Z}_p^d) \subset (\hat{k}^\times)^d$ is equivalent to the assertion that $a_{i,j} \in k$ for all (i,j).

Remark (7-2). In the case where $\dim A = 1$, A is a Tate curve. As is well-known, for any $q \in K^\times$ such that $v(q) \neq 0$, there is a Tate curve which is isomorphic to $K^\times/q^{\mathbb{Z}}$ as a rigid analytic group. So there are many elliptic curves whose p-divisible group does not admit a Hodge-Tate decomposition. In the higher dimensional case, the existence of an abelian variety corresponding to Λ is equivalent to the existence of a "polarization" on Λ (c.f. Mumford [8] Definition (1. 2)).

Next we consider the case (0-3-2). Let X be a p-divisible group over O_K which is an extension of $(\mathbb{Q}_p/\mathbb{Z}_p)^m$ by $(G_m(p))^n$, and let $\bar{X} = X \times_{O_K} \bar{K}$ be the reduction of X. Then the relation of the invariants of X and $\bar{X}$ defined after Lemma (6-3) is as follows:

$$Q(X) \equiv Q(\bar{X}) \bmod (U_K)^{mn},$$

where $U_K = \{x \in K^\times \mid v(1-x) > 0\}$. This follows from the commutative diagram

$$\begin{array}{ccc} H^1_{fl}(\operatorname{Spec} O_K, \mathbb{Z}/p^r\mathbb{Z}(1)) & \longrightarrow & H^1_{fl}(\bar{K}, \mathbb{Z}/p^r\mathbb{Z}(1)) \\ \wr\| & & \wr\| \\ O_K^\times/(O_K^\times)^{p^r} & \longrightarrow & \bar{K}^\times/(\bar{K}^\times)^{p^r} \\ \cup & & \cup \\ q_r & \longmapsto & \bar{q}_r \end{array}$$

We now consider A. Recall that

$$T_p(A) \otimes_{\mathbb{Z}_p} \mathbb{C}_p \cong \mathbb{C}_p^d \oplus \mathbb{C}_p(1)^d \quad \text{if and only if} \quad Q(A(p)) \in (\hat{k}^\times)^{d^2}.$$

As $\hat{k}^\times = \{1\}$, that $T_p(A)$ admits a Hodge-Tate decomposition shows that

$$(\mathscr{A} \times_{O_K} \bar{K})(p) \cong (\mathbb{Q}_p/\mathbb{Z}_p)^d \oplus (G_m(p))^d.$$

Thus the second assertion of Theorem 3 is proved. If $\bar{K}$ satisfies (0-4), we see by Proposition (6-1) that $\mathscr{A} \times_{O_K} \bar{K}$ is defined over $\bar{k}$. Using Proposition (6-2), we see that $\mathscr{A}$ is defined over O_k if and only if

$$Q(A(p)) \in (\hat{O}_k^\times)^{d^2} \subset (\hat{k}^\times)^{d^2}.$$

Thus we have shown the first assertion of Theorem 3. The third assertion is immediate from the following Lemma (7-3) and Remark (7-4).

Lemma (7-3). *For any* $\{q_{i,j}\}_{1 \leqq i,j \leqq d} \in (U_K)^{d^2} \subset (\hat{K}^\times)^{d^2}$, *there is a unique formal abelian scheme* $\mathscr{A}$ *over* O_K *such that*

$$(\mathscr{A} \times_{O_K} \bar{K})(p) \cong (Q_p/Z_p)^d \oplus (G_m(p))^d, \tag{7-3-1}$$

$$Q(A(p)) = \{q_{i,j}\}_{1 \leqq i,j \leqq d}. \tag{7-3-2}$$

This is immediate from Proposition (6-2) and Lemma (6-3).

Remark (7-4). In the case $d=1$, all $\mathscr{A}$ in Lemma (7-3) are algebraizable by [3] III Proposition 7. 2.

References

[1] *H. P. Epp*, Eliminating Wild Ramification, Invent. Math. **19** (1973), 235—249.

[2] *J.-M. Fontaine*, Formes Différentielles et Modules de Tate des Variétés Abéliennes sur les Corps Locaux, Invent. Math. **65** (1982), 379—409.

[3] *A. Grothendieck*, Revêtements Etales et Groupe Fondamental (SGA 1), Lecture Notes in Math. **224,** Berlin-Heidelberg-New York 1971.

[4] *K. Kato*, A generalization of local class field theory by using *K*-groups. II, J. Fac. Sci. Univ. Tokyo. IA, **27** (1980), 603—683.

[5] *J. Lubin, J.-P. Serre* and *J. Tate*, Elliptic curves and formal groups, Woods Hole Summer Institute 1964.

[6] *H. Miki*, On cyclic extensions of *p*-power degree over complete *p*-adic fields (in Japanese), Master's thesis, Tokyo 1973.

[7] *H. Miki*, On Z_p-extensions of complete *p*-adic power series fields and function fields, J. Fac. Sci. Univ. Tokyo. IA, **21** (1974), 377—393.

[8] *D. Mumford*, An Analytic Construction of Degenerating Abelian Varieties over Complete Rings, Compositio Math. **24** (1972), 239—272.

[9] *D. Mumford*, Geometric Invariant Theory, Berlin-Heidelberg-New York 1965.

[10] *M. Nagata*, Local rings, Interscience Tracts **13,** New York 1962.

[11] *M. Raynaud*, Modèles de Néron, C. R. Acad. Sci. Paris **262** (1966), 345—347.

[12] *M. Raynaud*, Variété Abéliennes et Géométrie Rigide, Actes Congrès intern. math. 1970 **1,** 473—477.

[13] *J.-P. Serre*, Corps Locaux, Paris 1962.

[14] *J.-P. Serre*, Sur Les Corps Locaux à Corps Résiduel Algébriquement Clos, Bull. Soc. Math. France **89** (1961), 105—154.

[15] *J. Tate*, *p*-Divisible Groups, Proc. of a conference on local fields, Berlin-Heidelberg-New York 1967, 158—183.

[16] *J. Tate*, Relations between K_2 and Galois Cohomology, Invent. Math. **36** (1976), 257—274.

Department of Mathematics, Faculty of Science, University of Tokyo, Hongo, Tokyo, Japan

Eingegangen 3. Januar 1985, in revidierter Form 24. April 1985

On analytic sets and functions with given isolated singularities

By *Wojciech Kucharz* at Albuquerque

1. Introduction

Let M be a paracompact connected m-dimensional real analytic manifold and let k be an integer, $1 \leqq k \leqq m-1$. Assume that there are given a discrete subset D of M and a family $\{V_x\}_{x \in D}$ such that V_x is an analytic set-germ at x of pure dimension k and x is its isolated singular point. Bochnak [2] and Bochnak and Risler [3] have proposed the following problem.

Problem. Does there exist a closed analytic subset W of M such that:

(1. 1) W is connected;

(1. 2) $W-D$ is an analytic submanifold of M;

(1. 3) for each x in D the germ of W at x is analytically equivalent to V_x?

We recall that two set-germs A and B at $x \in M$ are called C^r *equivalent* if there exists a local C^r diffeomorphism $\sigma : (M, x) \to (M, x)$ such that $B = \sigma(A)$, $r = 1, 2, \ldots, \infty, \omega$, (where ω stands for analytic).

The main result of this paper is the following.

Theorem 1. 4. *Given data M, D and $\{V_x\}$ as above, there exists a closed analytic subset W of M satisfying* (1. 1), (1. 2) *and* (1. 3) *in each of the following cases*:

(a) *M is noncompact*;

(b) $2k+1 \leqq m$;

(c) $k = m-1$;

(d) $m \leqq 4$;

(e) *V_x is coherent and is a complete intersection for each x in D.*

Moreover, if D is finite, then in all cases, except (a), *W can be chosen compact.*

Clearly, the same problem can be considered with real analytic data replaced by complex analytic ones. In fact, the first result of this kind has been obtained by Strehl [16] under the assumption that M is a complex Stein manifold and $k=m-1$. Bochnak [2] has relaxed the condition on V_x, assuming only that V_x is a complete intersection. He also has given a solution in the real case with each V_x being coherent and a complete intersection. However, in the real case, instead of (1. 3) he only has shown the C^r equivalence for r finite. Tang [17] has given a solution for $M=\mathbb{C}^m$ and $2k+1\leqq m$. In the real case, the problem has been solved by Bochnak et al. [5] under the assumption that each V_x is coherent and $k=m-1$.

We also have the following, closely related, result about constructing analytic functions with given isolated singularities.

Theorem 1. 5. *Let M be a connected noncompact real analytic manifold of dimension m, $m\geqq 2$. Assume that there are given a discrete subset D of M and a family $\{g_x\}_{x\in D}$, where $g_x:(M,x)\to\mathbb{R}$ is an analytic function-germ with x being its isolated critical point. Then there exists an analytic function $f: M\to\mathbb{R}$ such that f has no critical point outside D and for each x in D the germ of f at x is analytically equivalent to g_x.*

Here two function-germs $\varphi, \psi:(M,x)\to\mathbb{R}$ are called *analytically equivalent* if there exists a local analytical diffeomorphism $\sigma:(M,x)\to(M,x)$ such that $\psi=\varphi\circ\sigma$. Since every differentiable function on a compact manifold has at least two critical points, the assumption "M noncompact" is essential. Clearly, the assumption "$m\geqq 2$" is also necessary.

Theorem 1. 5 can be viewed as a generalization of Hirsch's result [9] that on every connected noncompact manifold there exists a differentiable function with no critical point.

2. A criterion of differential topology

We preserve the notation introduced in Section 1. Our goal is to give a criterion, in terms of differential topology, for the existence of a closed analytic subset W of M satisfying (1. 1), (1. 2) and (1. 3).

Given a C^r submanifold A of a C^r manifold Y (both A and Y with possibly nonempty boundary), $r=1, 2, \ldots, \infty, \omega$, we say that A is a neat submanifold if $\partial A=A\cap\partial Y$ and A is not tangent to ∂Y at any point $x\in\partial A$, i.e., $T_xA+T_x(\partial Y)=T_xY$. An embedding $f: X\to Y$ is called neat if $f(X)$ is a neat submanifold of Y. The set of all C^r maps from X to Y is denoted by $C^r(X, Y)$. All manifolds are assumed to be paracompact.

Let data M, D and $\{V_x\}_{x\in D}$, as in Section 1, be given. We fix families $\{U_x\}_{x\in D}$, $\{W_x\}_{x\in D}$ and $\{M_x\}_{x\in D}$ such that for each x in D:

(2. 1) U_x is a neighborhood of x in M, $U_x\cap U_{x'}=\emptyset$ for $x\neq x'$;

(2. 2) W_x is a closed analytic subset of U_x with a unique singular point at x such that the germ of W_x at x is equal to V_x;

(2. 3) M_x is an analytic submanifold with boundary of U_x which is diffeomorphic to the m-disc $D^m = \{z \in \mathbb{R}^m \mid \|z\| \leqq 1\}$, $x \in \operatorname{Int} M_x$ and ∂M_x is transversal to $W_x - \{x\}$.

Set

$$M_1 = M - \bigcup_{x \in D} \operatorname{Int} M_x \quad \text{and} \quad B = \bigcup_{x \in D} (\partial M_x \cap W_x).$$

Clearly, M_1 is an analytic submanifold with boundary of M and B is an analytic submanifold of ∂M_1.

The following criterion will play an important role.

Theorem 2. 4. *Assume that there exists a closed neat C^1 submanifold N of M_1 such that $\partial N = B$. Then there exists a closed analytic subset W of M satisfying* (1. 1), (1. 2) *and* (1. 3). *Moreover, if there is a compact N with the above properties, then W also can be chosen compact.*

In the proof of Theorem 2. 4, we shall be using the following two results which are well known but whose proofs are not readily available in the literature.

Lemma 2. 5. *Let A be an analytic neat submanifold of an analytic manifold Y. Then there exist a neighborhood U of ∂Y in Y and an analytic diffeomorphism $\varphi : U \to \partial Y \times [0, 1)$ such that $\varphi(x) = (x, 0)$ for x in ∂Y and $\varphi(A \cap U) = \partial A \times [0, 1)$.*

Lemma 2. 6. *Let X and Y be analytic manifolds with boundary and let $\varphi : \partial X \to \partial Y$ be an analytic map. Then the set $C^\omega(X, Y, \varphi)$ is dense in $C^1(X, Y, \varphi)$ in the Whitney C^1 topology, where*

$$C^r(X, Y, \varphi) = \{f \in C^r(X, Y) \mid f(x) = \varphi(x) \text{ for all } x \text{ in } \partial X\}$$

for $r = 1$ or ω.

The reader may consult [10] for proofs of the corresponding results in the C^∞ case and [11] for suggestions on how to obtain proofs in the analytic case.

We also need the following

Lemma 2. 7 ([10], p. 38 Corollary 1. 6 and p. 41 Exercise 11). *Let X and Y be C^1 manifolds with boundary. Let*

$$C^1(X, \partial X; Y, \partial Y) = \{f \in C^1(X, Y) \mid f(\partial X) \subset \partial Y\}.$$

Then the set

$$U = \{f \in C^1(X, \partial X; Y, \partial Y) \mid f \text{ is a closed neat embedding}\}$$

is open in $C^1(X, \partial X; Y, \partial Y)$ in the Whitney C^1 topology.

Proof of Theorem 2. 4. By cutting small holes in connected components of N and joining the remaining sets by k-dimensional "tubes" contained in $\operatorname{Int} M_1$, we may assume that N is connected. Note that the boundary ∂N of N is an analytic submanifold of ∂M_1. Consider N as an abstract C^1 manifold. One can find an analytic structure on N which is compatible with the C^1 structure and agrees with the analytic structure on ∂N. Denote the underlying topological space of N equipped with this analytic structure by N_0. The map $f : N_0 \to M_1$ defined by $f(x) = x$ for x in N_0 is of class C^1 and its restriction $f \mid \partial N_0 : \partial N_0 \to \partial M_1$ is analytic. By Lemmas 2. 6 and 2. 7

there exists an analytic map $f_0 : N_0 \to M_1$ which is a closed neat embedding and satisfies $f_0(x) = x$ for x in ∂N_0. Let us consider two couples of analytic manifolds with boundary (M_1, N_1) and (M_2, N_2), where

$$N_1 = f_0(N_0), \quad M_2 = \bigcup_{x \in D} M_x \quad \text{and} \quad N_2 = \bigcup_{x \in D} (M_x \cap (W_x - \{x\})).$$

Note that $\partial M_1 = \partial M_2$, $\partial N_1 = \partial N_2$ and N_i is a neat analytic submanifold with boundary of M_i, $i = 1, 2$. Now we shall carefully glue M_1 and M_2 along their common boundary. Let U_i be a neighborhood of ∂M_i in M_i such that there exists an analytic diffeomorphism $\varphi_i : U_i \to \partial M_i \times [0, 1)$ with $\varphi_i(x) = (x, 0)$ for x in ∂M_i and

$$\varphi_i(N_i \cap U_i) = \partial N_i \times [0, 1).$$

Denote by M' an analytic manifold, where $M' = M_1 \cup M_2$ as a set, every chart on $\operatorname{Int} M_1 \cup \operatorname{Int} M_2$ remains a chart on M' and the map $\varphi : U_1 \cup U_2 \to \partial M_1 \times (-1, 1)$ defined by $\varphi(x) = (\psi \circ \varphi_1)(x)$ for x in U_1 and $\varphi(x) = \varphi_2(x)$ for x in U_2, where

$$\psi : \partial M_1 \times [0, 1) \to \partial M_1 \times (-1, 0], \ \psi(y, t) = (y, -t),$$

is an analytic diffeomorphism. Note that

$$W' = N_1 \cup \bigcup_{x \in D} (M_x \cap W_x)$$

is a closed analytic subset of M'. By the uniqueness of gluing there is an analytic diffeomorphism $\sigma : M' \to M$. Let $\tau : M \to M$ be an analytic diffeomorphism satisfying $\tau(\sigma(x)) = x$ for x in D (cf. [4]). The set $W = \tau(\sigma(W'))$ is closed and analytic and satisfies (1. 1), (1. 2) and (1. 3).

The last part of Theorem 2. 4 is obvious. □

Theorem 2. 4 has the following interesting consequences.

Corollary 2. 8. *Given data M, D and $\{V_x\}_{x \in D}$, assume that there exists a closed analytic subset Z of M such that $Z - D$ is an analytic submanifold of M and for each x in D the germ Z_x of Z at x is C^1 equivalent to V_x. Then there exists a closed analytic subset W of M satisfying* (1. 1), (1. 2) *and* (1. 3). *Moreover, if there exists a compact Z with the above properties, then W also can be chosen compact.*

Proof. For each x in D let $\tau_x : (M, x) \to (M, x)$ be a local C^1 diffeomorphism such that $\tau_x(Z_x) = V_x$. Let $\gamma_x : (M, x) \to (M, x)$ be a local analytic diffeomorphism satisfying $\det d_x(\gamma_x \circ \tau_x) > 0$. Now we can find a C^1 diffeomorphism $\sigma : M \to M$ such that $\sigma(x) = x$ and the germ of σ at x is equal to $\gamma_x \circ \tau_x$ for all x in D. Clearly, there exists a family $\{U_x\}_{x \in D}$ of open sets in M such that $x \in U_x$, $U_x \cap U_{x'} = \emptyset$ for $x \neq x'$, and $W_x = U_x \cap \sigma(Z)$ is an analytic subset of U_x for which x is a unique singular point and whose germ at x is equal to $\gamma_x(V_x)$. Let M_x be an analytic submanifold with boundary of U_x such that M_x is diffeomorphic to the m-disc D^m, $x \in \operatorname{Int} M_x$ and ∂M_x is transversal to $W_x - \{x\}$. Note that

$$N = \sigma(Z) - \bigcup_{x \in D} \operatorname{Int} M_x$$

is a closed neat C^1 submanifold of

$$M - \bigcup_{x \in D} \operatorname{Int} M_x$$

with

$$\partial N = \bigcup_{x \in D} (\partial M_x \cap W_x).$$

Hence, by Theorem 2. 4, there exists a closed analytic subset W of M such that W is connected, $W-D$ is an analytic submanifold of M and for each x in D, the germ of W at x is analytically equivalent to $\gamma_x(V_x)$. Since γ_x is a local analytic diffeomorphism, the conclusion follows. The last part of Corollary 2. 8 is obvious. □

Corollary 2. 9. *Let data M, D and $\{V_x\}_{x\in D}$ be given. Assume that for each x in D there exist a local analytic diffeomorphism $\sigma_x: (M, x) \to (\mathbb{R}^m, 0)$ and a compact analytic subset Z_x of $\mathbb{R}^m$ such that $0 \in \mathbb{R}^m$ is a unique singular point of Z_x and $\sigma_x(V_x)$ is equal to the germ of Z_x at $0 \in \mathbb{R}^m$. Then there exists a closed analytic subset W of M satisfying* (1. 1), (1. 2) *and* (1. 3). *Moreover, if D if finite, then W can be chosen compact.*

Proof. We can find two families $\{U_x\}_{x\in D}$ and $\{\tau_x\}_{x\in D}$ such that U_x is a neighborhood of x in M, $U_x \cap U_{x'} = \emptyset$ for $x \neq x'$ and $\tau_x: U_x \to \mathbb{R}^m$ is an analytic diffeomorphism. Setting

$$Z = \bigcup_{x\in D} \tau_x^{-1}(Z_x)$$

the conclusion follows from Corollary 2. 8. □

3. Analytic sets with isolated singularities

We are ready to prove the main result of this paper.

Proof of Theorem 1. 4. (a) Let $\{U_x\}_{x\in D}$, $\{W_x\}_{x\in D}$ and $\{M_x\}_{x\in D}$ be families satisfying (2. 1), (2. 2) and (2. 3) and let $\varphi_x: \operatorname{Int} M_x \to \mathbb{R}^m$ be an analytic diffeomorphism. Denote by S^m the unit m-sphere and by $\rho: S^m - \{a\} \to \mathbb{R}^m$ the stereographic projection from $a = (0, \ldots, 0, 1)$. Let b_x be a point in

$$S^m - \left(\rho^{-1}(\varphi_x(W_x \cap \operatorname{Int} M_x)) \cup \{a, -a\}\right).$$

We choose an analytic diffeomorphism $\psi_x: S^m - \{b_x\} \to \operatorname{Int} M_x$ with $\psi_x(-a) = x$. Note that

$$W_x' = (\psi_x \circ \rho^{-1} \circ \varphi_x)(W_x \cap \operatorname{Int} M_x)$$

is a closed analytic subset of $\operatorname{Int} M_x$ with two distinct singular point x and $\psi_x(a)$ and the germ of W_x' at x is analytically equivalent to V_x. Let M_x' be an analytic submanifold with boundary of $\operatorname{Int} M_x$ such that M_x' is diffeomorphic to D^m, $x \in \operatorname{Int} M_x'$, $\psi_x(a) \notin M_x'$ and $\partial M_x'$ is transversal to $W_x' - \{x, \psi_x(a)\}$. Now we can find (cf. [9]) two families $\{H_x\}_{x\in D}$ and $\{G_x\}_{x\in D}$ such that H_x is a closed C^1 submanifold with boundary of M diffeomorphic to the halfspace $\mathbb{R}^{m-1} \times [0, \infty)$, $\psi_x(a) \in \operatorname{Int} H_x$, G_x is an open neighborhood of H_x, $G_x \cap G_{x'} = \emptyset$ for $x \neq x'$ and $\bigcup_{x\in D} M_x'$ is disjoint from $G = \bigcup_{x\in D} G_x$. Set

$$U = M - \bigcup_{x\in D} H_x \quad \text{and} \quad U_1 = U - \bigcup_{x\in D} \operatorname{Int} M_x'.$$

Note that

$$N = \bigcup_{x\in D} \left((W_x' \cap U) - \operatorname{Int} M_x'\right)$$

is a closed neat analytic submanifold with boundary of U_1 and

$$\partial N = \bigcup_{x\in D} (W_x' \cap \partial M_x').$$

By Theorem 2. 4 there exists a closed analytic subset Z' of U such that $Z'-D$ is an analytic submanifold of U and the germ of Z' at x is analytically equivalent to V_x for each x in D. By [9], Lemma 3. 6, there exists a C^1 diffeomorphism $g: M \to U$ such that $g(y)=y$ for y in $M-G$. Therefore, we also can find an analytic diffeomorphism $h: M \to U$ such that $h(x)=x$ for x in D. Now it suffices to set $Z=h^{-1}(Z')$ and apply Corollary 2. 8.

(b) Part (b) follows from Corollary 2. 9 and the following lemma.

Lemma 3. 1. *Let U be a neighborhood of 0 in $\mathbb{R}^m$ and let V be a closed analytic subset of U of pure dimension k with 0 being its unique singular point. Assume that $2k+1 \leqq m$. Then there exists a compact analytic subset W of $\mathbb{R}^m$ such that 0 is its unique singular point and the germs of V and W at 0 are analytically equivalent.*

Proof. By Hironaka's desingularization theorem [8], there exist an analytic manifold X and a proper analytic map $\pi: X \to V$ such that the restriction

$$\pi | X-\pi^{-1}(0): X-\pi^{-1}(0) \to V-\{0\}$$

is an analytic diffeomorphism. For $\varepsilon>0$ set $D_\varepsilon=\{x \in \mathbb{R}^m \mid \|x\| \leqq \varepsilon\}$. If ε is sufficiently small, then $D_\varepsilon \subset U$ and ∂D_ε is transversal to $V-\{0\}$. Note that $N=\pi^{-1}(D_\varepsilon)$ is a compact analytic manifold with boundary and ∂N is diffeomorphic to $\partial D_\varepsilon \cap V$. Since $2k+1 \leqq m$, there exists a closed neat C^∞ embedding

$$\varphi: N \to \mathbb{R}^m - \operatorname{Int} D_\varepsilon$$

such that $\partial(\varphi(N))=\partial D_\varepsilon \cap V$. Now the conclusion follows from Theorem 2. 4.

(c) To prove (c), we only need to apply Theorem 2. 4, Corollary 2. 9 and the following lemma.

Lemma 3. 2. *Let A be a closed boundaryless C^∞ submanifold of ∂D^m, where*

$$D^m=\{x \in \mathbb{R}^m \mid \|x\| \leqq 1\}.$$

If $\dim A=m-2$, then there exists a compact neat C^∞ submanifold X of $\mathbb{R}^m-\operatorname{Int} D^m$ such that $\partial X=A$.

Proof. Since A bounds in ∂D^m, there exists a C^∞ function $f: \partial D^m \to \mathbb{R}$ with $0 \in \mathbb{R}$ being its regular value and $A=f^{-1}(0)$. Clearly,

$$Y=\{(x, t) \in \partial D^m \times [0, \infty) \mid f(x)=t\}$$

is a compact neat C^∞ submanifold of $\partial D^m \times [0, \infty)$ with $\partial Y=A \times \{0\}$. Note that

$$\varphi: \partial D^m \times [0, \infty) \to \mathbb{R}^m - \operatorname{Int} D^m$$

defined by $\varphi(x, t)=(1+t)\,x$ is a C^∞ diffeomorphism. It suffices to set $X=\varphi(Y)$.

(d) If $m \leqq 3$ or $m=4$ and $k=1$ or 3, then the conclusion follows from (b) and (c), respectively. If $m=4$ and $k=2$, it suffices to apply Theorem 2. 4 and Corollary 2. 9 and use the fact that any compact one-dimensional boundaryless C^∞ submanifold of ∂D^4 is the boundary of a compact neat C^∞ submanifold of $\mathbb{R}^4-\operatorname{Int} D^4$.

(e) It suffices to show the following lemma.

Lemma 3. 3. *Let V be a germ at $0 \in \mathbb{R}^m$ of a coherent analytic set with isolated singular point at 0. Assume that V is a complete intersection. Then there exists a compact analytic subset W of $\mathbb{R}^m$ such that $0 \in \mathbb{R}^m$ is the unique singular point of W and the germ of W at 0 is analytically equivalent to V.*

Proof. Let $p = \operatorname{codim} V$. By the assumption, there are an open neighborhood U of 0 in $\mathbb{R}^m$ and an analytic map $f = (f_1, \ldots, f_p) : U \to \mathbb{R}^p$ such that the germ of $f^{-1}(0)$ at 0 is equal to V and $0 \in \mathbb{R}^p$ is a regular value of the restriction of f to $U - \{0\}$. Let φ be the sum of squares of f_i's and all $p \times p$ minors of the matrix $\left(\frac{\partial f_i}{\partial x_j}\right)$, $i = 1, \ldots, p$, $j = 1, \ldots, m$. Observe that, in particular, $\varphi^{-1}(0) = \{0\}$. Fix a positive real number ε such that $D_\varepsilon = \{x \in \mathbb{R}^m \mid \|x\| < \varepsilon\}$ is contained in U. Consider the map $g = (g_1, \ldots, g_p)$ from U to $\mathbb{R}^p$, where $g_i = f_i + t_i \varphi^2$ for some $t_i \in \mathbb{R}$, $i = 1, \ldots, p$. By a transversality argument, one can find arbitrarily large t_i's such that $0 \in \mathbb{R}^p$ is a regular value of the restriction of g to $U - \{0\}$. Hence $0 \in \mathbb{R}^m$ is the unique singular point of $g^{-1}(0)$ and if t_i's are large enough, the connected component W of $g^{-1}(0)$ containing 0 is contained in $\operatorname{Int} D_\varepsilon$. Finally, by [18], the germ of W at 0 is analytically equivalent to V.

The last part of Theorem 1. 4 is obvious. □

4. Analytic functions with isolated singularities

This section is devoted to proving Theorem 1. 5.

Proof of Theorem 1. 5. One can find a C^∞ function $g : M \to \mathbb{R}$ such that the germ of g at x is equal to g_x for all x in D. Pick a neighborhood U of D such that g has no critical point on $U - D$. Let U_1, U_2 be neighborhoods of D in M such that $Cl\, U_2 \subset U_1$ and $Cl\, U_1 \subset U$, where $Cl\, U_i$ denotes the closure of U_i, $i = 1, 2$. Denote by $J^1(M, \mathbb{R})$ the space of 1-jets from M to $\mathbb{R}$, by $\pi : J^1(M, \mathbb{R}) \to M$ the canonical projection, and by $j^1\varphi$ the 1-jet of a map $\varphi : M \to \mathbb{R}$. We define the submanifold W of $J^1(M, \mathbb{R})$ as follows

$$W = \{j^1_x \psi \in J^1(M, \mathbb{R}) \mid x \in M - Cl\, U_1,\ d_x \psi = 0\}.$$

By [6], Corollary 4. 11, there exists a C^∞ map $h : M \to \mathbb{R}$, arbitrarily close to g in the Whitney C^∞ topology, such that $h = g$ on U_2 and $j^1 h$ is transversal to W. Clearly, h has only nondegenerate (thus isolated) critical points on $M - Cl\, U_1$. By taking h sufficiently close to g, we may assume that h has no critical point on $Cl\, U_1 - U_2$. Therefore h has only isolated critical points. Now, as in the proof of part (a) of Theorem 1. 4, we can find an open subset M_0 of M and an analytic diffeomorphism $\varphi : M \to M_0$ such that $D \subset M_0$, h has no critical point on $M_0 - D$ and $\varphi(x) = x$ for x in D. Note that the function $h \circ \varphi$ is analytic in a neighborhood of D and has no critical point outside D. The conclusion follows from [15], Corollary 6. 6 and [1]. □

References

[1] *M. Artin*, On the solution of analytic equations, Invent. math. **5** (1968), 277—291.
[2] *J. Bochnak*, Sur les ensembles analytiques à singularités données a priori, Math. Ann. **216** (1975), 269—271.
[3] *J. Bochnak, J. J. Risler*, Analyse différentielle et géométrie analytique, Quelques questions ouvertes, Lecture Notes in Math. **535**, Berlin-Heidelberg-New York 1976, 63—69.
[4] *J. Bochnak, W. Kucharz, M. Shiota*, On equivalence of ideals of real global analytic functions and the 17th Hilbert problem, Invent. math. **63** (1981), 403—421.
[5] *J. Bochnak, W. Kucharz, M. Shiota*, On algebraicity of global real analytic sets and functions, Invent. math. **70** (1983), 115—156.
[6] *M. Golubitsky, V. Guillemin*, Stable mappings and their singularities, Berlin-Heidelberg-New York 1973.
[7] *H. Grauert*, On Levi's problem and the imbedding of real analytic manifolds, Ann. Math. **68** (1958), 460—472.
[8] *H. Hironaka*, Resolution of singularities of an algebraic variety over a field of characteristic zero. I—II, Ann. Math. **79** (1964), 109—326.
[9] *M. W. Hirsch*, On imbedding differentiable manifolds in Euclidean space, Ann. Math. **73** (1961), 566—571.
[10] *M. W. Hirsch*, Differential topology, Berlin-Heidelberg-New York 1976.
[11] U. Hirsch, Some remarks on analytic foliations and analytic branched coverings, Math. Ann. **248** (1980), 139—152.
[12] *D. Husemoller*, Fiber bundles, Berlin-Heidelberg-New York 1975, second edition.
[13] *J. Milnor*, Singular points of complex hypersurfaces. Ann. Math. Studies **61**, Princeton 1968.
[14] *R. Narasimhan*, Analysis on real and complex manifolds, Amsterdam-London 1968.
[15] *M. Shiota*, Equivalence of differentiable mappings and analytic mappings, Publ. Math. I.H.E.S. **54** (1981), 38—122.
[16] *B. Strehl*, Analytische Hyperflächen mit vorgegebenen Singularitäten, Math. Ann. **200** (1973), 165—173.
[17] *C. T. Tang*, Analytic subvarieties in $\mathbb{C}^n$ with given isolated singularities, Math. Ann. **229** (1977), 47—51.
[18] *J. Cl. Tougeron*, Idéaux de fonctions différentiables. I, Ann. Inst. Fourier **18** (1969), 177—240.

Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico 87131, U.S.A.

Eingegangen 7. Mai 1985

On Waring's problem for cubes

By *R. C. Vaughan* at London

1. Introduction

Let $\mathscr{R}_s(N)$ denote the number of representations of N as the sum of s positive cubes and let

$$\mathfrak{S}_s(N)=\sum_{q=1}^{\infty}\sum_{\substack{a=1\\(a,q)=1}}^{q}(S(q,a)q^{-1})^s e\left(-\frac{aN}{q}\right) \tag{1.1}$$

with

$$S(q,a)=\sum_{r=1}^{q} e\left(\frac{ar^3}{q}\right) \tag{1.2}$$

denote the corresponding singular series. The asymptotic formula

$$\mathscr{R}_s(N)\sim\frac{\Gamma\left(\frac{4}{3}\right)^s}{\Gamma\left(\frac{s}{3}\right)}\mathfrak{S}_s(N)N^{\frac{1}{3}s-1} \tag{1.3}$$

is known when $s\geqq 9$ (see, for example, Chapter 2 of Vaughan [V1]), conjectured when $4\leqq s\leqq 8$, and established conditionally by Hooley [H3] when $s=7$ or 8, the condition being the assumption of the Riemann hypothesis for a certain rather recondite L-function.

In our first theorem we establish (1.3) when $s=8$.

Theorem 1. *We have*

$$\mathscr{R}_8(N)=\frac{\Gamma\left(\frac{4}{3}\right)^8}{\Gamma\left(\frac{8}{3}\right)}\mathfrak{S}_8(N)N^{\frac{5}{3}}+O\left(N^{\frac{5}{3}}(\log N)^{c-1+\varepsilon}\right)$$

where c is a constant with $0\leqq c\leqq\frac{4}{\pi}-1$.

Throughout ε is a positive number, not necessarily the same in different formulae, large means greater than a number depending on ε, and implicit constants may depend on ε.

Our method enables us to establish two further theorems of some interest.

Theorem 2. *Let $R(P)$ denote the number of solutions of*

$$x_1^3 + \cdots + x_4^3 = x_5^3 + \cdots + x_8^3$$

with $x_i \leqq P$. Then

$$R(P) = J\mathfrak{S}P^5 + O\left(P^5(\log P)^{c-1+\varepsilon}\right)$$

where

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-8}|S(q,a)|^8,$$

$$J = \int_{-\infty}^{\infty} \left|\int_0^1 e(\beta x^3)dx\right|^8 d\beta$$

and c is a constant with $0 \leqq c \leqq \frac{4}{\pi} - 1$.

Previously Hooley [H1], § 4 had obtained (implicitly) the estimate

$$R(P) \ll P^5(\log P)^{\frac{4}{\pi}-1+\varepsilon}.$$

Theorem 3. *We have*

$$\sum_{n=1}^{N} \left(\mathscr{R}_4(n) - \Gamma\left(\frac{4}{3}\right)^3 \mathfrak{S}_4(n)n^{\frac{1}{3}}\right)^2 \ll N^{\frac{5}{3}}(\log N)^{c-1+\varepsilon}$$

where c is a constant with $0 \leqq c \leqq \frac{4}{\pi} - 1$.

Davenport [D1] has shown that the number $E(N)$ of natural numbers not exceeding N which are not the sum of four positive cubes satisfies

$$E(N) \ll N^{\frac{29}{30}+\varepsilon}. \tag{1.4}$$

By a variant of our methods we are able to improve upon this substantially.

Theorem 4. *We have $E(N) \ll N^{\frac{103}{115}+\varepsilon}$.*

Linnik [L] has shown that every large natural number is the sum of seven positive cubes, i.e. $G(3) \leqq 7$, and Watson [W] has given a very elegant proof of this. Neither the argument of Linnik nor that of Watson gives a very good bound for $\mathscr{R}_7(N)$. Although we are not able to obtain the asymptotic formula (1.3) when $s = 7$ we are able to obtain a good lower bound for $\mathscr{R}_7(N)$.

Theorem 5. *Suppose that N is sufficiently large. Then*

$$\mathscr{R}_7(N) \gg N^{\frac{142}{115}}(\log N)^{-5}.$$

This is within $N^{\frac{34}{345}}(\log N)^5$ of this conjectured order of magnitude given by (1. 3). Hitherto the Hardy-Littlewood method has been able only to establish $G(3) \leqq 8$ (Davenport [D1]).

By adapting certain ideas from Hooley [H3] it is very likely that the exponents $\frac{103}{115}$ and $\frac{142}{115}$ in Theorems 4 and 5 could be replaced by $\frac{131}{147}$ and $\frac{26}{21}$ respectively.

Let $\mathcal{N}(N)$ denote the number of natural numbers not exceeding N which are the sum of three cubes of natural numbers. In Vaughan [V3] it was shown that

$$\mathcal{N}(N) \gg N^{\frac{8}{9}-\varepsilon} \tag{1. 5}$$

improving earlier results of Davenport [D1], [D2]. As a by-product of our method we obtain a further improvement on this.

Theorem 6. *Suppose that* $N \geqq 3$. *Then*

$$\mathcal{N}(N) \gg N^{\frac{19}{21}-\varepsilon}.$$

We base our proof of Theorem 6 on the following theorem, which we establish in § 2.

Theorem A. *Suppose that* $M \leqq P^{\frac{1}{3}}$ *and* $Q = \frac{P}{M}$, *and let* S *denote the number of solutions of*

$$x_1^3 + m^3(y_1^3 + y_2^3) = x_2^3 + m^3(y_3^3 + y_4^3) \tag{1. 6}$$

subject to

$$x_i \leqq P, \quad y_i \leqq Q, \quad M < m \leqq 2M, \tag{1. 7}$$

$$(m, x_i) = 1. \tag{1. 8}$$

Then

$$S \ll MP^{1+\varepsilon}Q^2(P^{\frac{1}{2}}M^{-\frac{7}{2}} + 1 + M^2P^{-\frac{1}{2}}).$$

The deduction of Theorem 6 is standard and described briefly in § 3.

To establish Theorems 1, 2 and 3 we require a refinement of Theorem A. This enables us to establish

Theorem B. *Let* $\mathfrak{m}$ *denote the set of* α *in* $(0, 1]$ *with the property that whenever* $\left|\alpha - \frac{a}{q}\right| \leqq q^{-1}P^{-\frac{9}{4}}$ *with* $(a, q) = 1$ *we have* $q > P^{\frac{3}{4}}$. *Further let*

$$f(\alpha) = \sum_{x \leqq P} e(\alpha x^3).$$

Then there is a constant c with $0 \leqq c \leqq \frac{4}{\pi} - 1$ such that

$$\int_{\mathfrak{m}} |f(\alpha)|^8 \, d\alpha \ll P^5 (\log P)^{c-1+\varepsilon}.$$

The deduction of Theorems 1, 2 and 3 from Theorem B is via standard applications of the Hardy-Littlewood method (see Chapters 2, 3 and 4 of Vaughan [V1], for example) and so we suppress the details. The refinements that are required to the proof of Theorem A and the deduction of Theorem B are given in § 4.

In §§ 5—7 we establish a technical lemma which, in a certain sense, is an iterated version of Theorem A. The technical lemma is then converted into a suitable minor arc estimate in § 8. We then prove Theorem 4 in § 9 and Theorem 5 in § 10.

2. The proof of Theorem A

For a given m, let $\mathscr{A}(u)$ denote the set of solutions of

$$z^3 \equiv u \pmod{m^3}.$$

Then

$$\text{card}\, \mathscr{A}(u) \ll m^{\varepsilon} \quad ((u, m) = 1) \tag{2.1}$$

as can readily be verified by first of all considering the special case when m is a power of a prime.

Let

$$f_m(\alpha, z) = \sum_{\substack{x \leqq P \\ x \equiv z \pmod{m^3}}} e(\alpha x^3)$$

and

$$g(\beta) = \sum_{y \leqq Q} e(\beta y^3). \tag{2.2}$$

Clearly in each solution of (1.6) we have $x_1^3 \equiv x_2^3 \pmod{m^3}$. Thus

$$S = \sum_{M < m \leqq 2M} S_m$$

where

$$S_m = \int_0^1 F_m(\alpha) \, |g(m^3 \alpha)|^4 \, d\alpha$$

and

$$F_m(\alpha) = \sum_{\substack{u=1 \\ (u,m)=1}}^{m^3} \Big| \sum_{z \in \mathscr{A}(u)} f_m(\alpha, z) \Big|^2.$$

By Cauchy's inequality and (2.1)

$$F_m(\alpha) \ll m^{\varepsilon} \sum_{\substack{u=1 \\ (u,m)=1}}^{m^3} \sum_{z \in \mathscr{A}(u)} |f_m(\alpha, z)|^2 = m^{\varepsilon} \sum_{\substack{z=1 \\ (z,m)=1}}^{m^3} |f_m(\alpha, z)|^2.$$

Therefore

$$S \ll M^{\varepsilon} S_0$$

where S_0 is the number of solutions of (1. 6), (1. 7), (1. 8) subject to

$$x_1 \equiv x_2 \pmod{m^3}. \tag{2. 3}$$

Since the number of solutions of $y_1^3 + y_2^3 = n$ is $O(n^{\varepsilon})$, it follows that the terms in (1. 6) with $x_1 = x_2$ contribute

$$\ll MPQ^{2+\varepsilon}$$

to S_0. Thus it suffices to show that

$$S_1 \ll MP^{1+\varepsilon}Q^2(P^{\frac{1}{2}}M^{-\frac{7}{2}} + 1 + M^2P^{-\frac{1}{2}}) \tag{2. 4}$$

where S_1 is the number of solutions of (1. 6), (1. 7), (1. 8), (2. 3) with $x_2 > x_1$.

On writing $h = \dfrac{x_2 - x_1}{m^3}$, (1. 2) becomes

$$h(3(2x_1 + hm^3)^2 + h^2m^6) = 4(y_1^3 + y_2^3 - y_3^3 - y_4^3).$$

Therefore it now suffices to show that

$$S_2 \ll MP^{1+\varepsilon}Q^2(P^{\frac{1}{2}}M^{-\frac{7}{2}} + 1 + M^2P^{-\frac{1}{2}}) \tag{2. 5}$$

where S_2 is the number of solutions of

$$h(3x^2 + h^2m^6) = 4(y_1^3 + y_2^3 - y_3^3 - y_4^3) \tag{2. 6}$$

with

$$x \leqq 2P, \quad h \leqq PM^{-3}, \quad M < m \leqq 2M \tag{2. 7}$$

and the y_i satisfying (1. 7).

For brevity write

$$H = PM^{-3} \tag{2. 8}$$

and define

$$F(\beta) = \sum_{x \leqq 2P} e(\beta x^2), \tag{2. 9}$$

$$G(\beta) = \sum_{M < m \leqq 2M} e(\beta m^6). \tag{2. 10}$$

Then

$$S_2 = \int_0^1 \sum_{h \leqq H} F(3\alpha h)\, G(\alpha h^3)\, |g(4\alpha)|^4\, d\alpha. \tag{2. 11}$$

Let $\mathfrak{M}(q, a)$ denote the interval

$$\left\{\alpha: \left|\alpha - \frac{a}{q}\right| \leqq (qPH)^{-1}\right\}$$

and let $\mathfrak{U} = (P^{-1}H^{-1}, 1 + P^{-1}H^{-1}]$. We may suppose that P is large. Then the $\mathfrak{M}(q, a)$ with $1 \leqq a \leqq q \leqq P$ and $(a, q) = 1$ are disjoint and contained in $\mathfrak{U}$. Let $\mathfrak{M}$ denote their union, and let $\mathfrak{m} = \mathfrak{U} \setminus \mathfrak{M}$. Then, by (2. 11),

$$S_2 = I_{\mathfrak{M}} + I_{\mathfrak{m}} \tag{2. 12}$$

where for $\mathscr{B} = \mathfrak{M}$ or $\mathfrak{m}$,

$$I_{\mathscr{B}} = \int_{\mathscr{B}} \sum_{h \leqq H} F(3\alpha h)\, G(\alpha h^3)\, |g(4\alpha)|^4\, d\alpha. \tag{2. 13}$$

By an argument similar to that of the proof of the Lemma in [V3] we have

$$\sum_{h \leqq H} |F(3\alpha h)|^2 \ll (HP^2 q^{-1} + HP + q) P^{\varepsilon} \tag{2. 14}$$

whenever $\left|\alpha - \frac{a}{q}\right| \leqq q^{-2}$ and $(a, q) = 1$. Following the pattern established in [V3] we find that for $\alpha \in \mathfrak{M}(q, a)$ we have

$$g(4\alpha) \ll q^{-\frac{1}{3}} Q \left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{-\frac{1}{3}} + P^{\frac{1}{2}+\varepsilon}$$

and

$$\sum_{h \leqq H} F(3\alpha h)\, G(\alpha h^3) \ll \sum_{h \leqq H} |F(3\alpha h)|\, M \ll HMP^{1+\varepsilon} q^{-\frac{1}{2}}.$$

Therefore, by (2. 13),

$$\begin{aligned} I_{\mathfrak{M}} &\ll \sum_{q \leqq P} HMP^{1+\varepsilon} q^{-\frac{1}{2}} \left(\int_0^{\infty} \frac{Q^4 q^{-\frac{4}{3}}}{(1 + Q^3\beta)^{\frac{4}{3}}}\, d\beta + P^{2+\varepsilon} q^{-1} H^{-1} P^{-1} \right) \\ &\ll HMP^{1+\varepsilon} Q P^{\frac{1}{6}} + MP^{2+2\varepsilon} P^{\frac{1}{2}}. \end{aligned}$$

Hence, by (2. 12),

$$S_2 = I_{\mathfrak{m}} + O\left(MP^{1+\varepsilon} Q^2 (P^{\frac{1}{6}} M^{-2} + M^2 P^{-\frac{1}{2}})\right). \tag{2. 15}$$

By (2. 13) and the Cauchy-Schwarz inequality we have

$$I_{\mathfrak{m}} \ll (IJ)^{\frac{1}{2}} \tag{2. 16}$$

where

$$I = \int_{\mathfrak{m}} \sum_{h \leqq H} |F(3\alpha h)|^2\, |g(4\alpha)|^4\, d\alpha \tag{2. 17}$$

and

$$J = \int_0^1 \sum_{h \leqq H} |G(\alpha h^3)|^2\, |g(4\alpha)|^4\, d\alpha. \tag{2. 18}$$

It is readily seen that for each $\alpha \in \mathfrak{m}$ there exist a, q with $\left|\alpha - \frac{a}{q}\right| \leqq (qHP)^{-1}$, $P < q \leqq HP$ and $(a, q) = 1$. Hence, by (2. 14),

$$\sum_{h \leqq H} |F(3\alpha h)|^2 \ll HP^{1+\varepsilon}.$$

Therefore, by Hua's Lemma (Lemma 2. 5 of Vaughan [V1]) we have

$$I \ll HP^{1+\varepsilon}Q^{2+\varepsilon}.$$

Thus, by (2. 16),

$$I_{\mathfrak{m}} \ll H^{\frac{1}{2}}P^{\frac{1}{2}+\varepsilon}QJ^{\frac{1}{2}}. \tag{2. 19}$$

By (2. 10),

$$\sum_{h \leqq H} |G(\alpha h^3)|^2 = \sum_{h \leqq H} \sum_{M < m_1 \leqq 2M} \sum_{M < m_2 \leqq 2M} e(\alpha h^3 (m_2^6 - m_1^6))$$

$$\ll HM + \sum_{M < m_1 \leqq 2M} \sum_{M < m_2 \leqq 2M} |E(\alpha(m_2^6 - m_1^6))|$$

where

$$E(\beta) = \sum_{h \leqq H} e(\beta h^3).$$

Therefore, by (2. 18) and Hua's Lemma we have

$$J \ll HMQ^{2+\varepsilon} + \sum_{M < m_1 \leqq 2M} \sum_{M < m_2 \leqq 2M} K(m_2^6 - m_1^6) \tag{2. 20}$$

where

$$K(d) = \int_0^1 |E(\alpha d)|\, |g(4\alpha)|^4 \, d\alpha.$$

By the change of variable $\beta = \alpha d$ we have

$$K(d) = d^{-1} \int_0^d |E(\beta)| \left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta. \tag{2. 21}$$

Let $\mathfrak{N}(q, a)$ denote the interval

$$\left\{\alpha : \left|\alpha - \frac{a}{q}\right| \leqq (6qH^2)^{-1}\right\}$$

and let $\mathfrak{B} = ((6H^2)^{-1}, d + (6H^2)^{-1}]$. Then the $\mathfrak{N}(q, a)$ with $1 \leqq a \leqq dq$, $q \leqq H$ and $(a, q) = 1$ are disjoint and contained in $\mathfrak{B}$. Let $\mathfrak{N}$ denote their union and let $\mathfrak{n} = \mathfrak{B} \setminus \mathfrak{N}$.

Whenever $\beta \in \mathfrak{n}$ there exist a, q with $1 \leqq a \leqq dq$, $H < q \leqq 6H$, $(a, q) = 1$ and $\left|\alpha - \frac{a}{q}\right| \leqq q^{-2}$. Hence, by Weyl's inequality (Lemma 2. 4 of Vaughan [V1]) we have

$$E(\beta) \ll H^{\frac{3}{4}+\varepsilon}.$$

Therefore, by Hua's Lemma,

$$d^{-1}\int_{\mathfrak{n}}|E(\beta)|\left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta \ll H^{\frac{3}{4}+\varepsilon}d^{-1}\int_0^d\left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta$$
$$\ll H^{\frac{3}{4}+\varepsilon}Q^{2+\varepsilon}.$$

Thus, by (2. 21),

$$K(d)=d^{-1}\int_{\mathfrak{R}}|E(\beta)|\left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta+O(H^{\frac{3}{4}+\varepsilon}Q^{2+\varepsilon}). \tag{2. 22}$$

By Hölder's inequality

$$d^{-1}\int_{\mathfrak{R}}|E(\beta)|\left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta\leqq L(d)^{\frac{1}{4}}M_4(d)^{\frac{1}{2}}M_8(d)^{\frac{1}{4}} \tag{2. 23}$$

where

$$L(d)=d^{-1}\int_{\mathfrak{R}}|E(\beta)|^4 d\beta$$

and

$$M_k(d)=d^{-1}\int_0^d\left|g\left(\frac{4\beta}{d}\right)\right|^k d\beta.$$

By a straightforward adaptation of the methods of § 4. 4 of Vaughan [V1] we have

$$L(d)\ll H^{1+\varepsilon}.$$

Moreover, by the change of variable $\alpha=\frac{4\beta}{d}$,

$$M_k(d)=4\int_0^1|g(\alpha)|^k d\alpha.$$

Hence, by Lemma 2. 5 of Vaughan [V1], (2. 22) and (2. 23),

$$K(d)\ll H^{\frac{1}{4}}Q^{\frac{9}{4}}P^{\varepsilon}+H^{\frac{3}{4}}Q^2P^{\varepsilon}.$$

Thus, by (2. 19) and (2. 20),

$$I_{\mathfrak{m}}\ll H^{\frac{1}{2}}P^{\frac{1}{2}+\varepsilon}Q(H^{\frac{1}{2}}M^{\frac{1}{2}}Q+MH^{\frac{1}{8}}Q^{\frac{9}{8}}+MH^{\frac{3}{8}}Q)$$
$$=MP^{1+\varepsilon}Q^2(P^{\frac{1}{2}}M^{-\frac{7}{2}}+P^{\frac{1}{4}}M^{-2}+P^{\frac{3}{8}}M^{-\frac{21}{8}}).$$

Therefore, by (2. 15)

$$S_2(MP^{1+\varepsilon}Q^2)^{-1}\ll P^{\frac{1}{2}}M^{-\frac{7}{2}}+P^{\frac{1}{4}}M^{-2}+P^{\frac{3}{8}}M^{-\frac{21}{8}}+M^2P^{-\frac{1}{2}}.$$

When $M\leqq P^{\frac{1}{7}}$ this is $\ll P^{\frac{1}{2}}M^{-\frac{7}{2}}$, when $P^{\frac{1}{7}}<M\leqq P^{\frac{1}{4}}$ it is $\ll 1$ and when $M>P^{\frac{1}{4}}$ it is $\ll M^2P^{-\frac{1}{2}}$. This gives (2. 5) and so completes the proof of the theorem.

3. The proof of Theorem 6

Let $M=P^{\frac{1}{7}}$ and let $r(n)$ denote the number of solutions of

$$x_1^3+m^3(y_1^3+y_2^3)=n \tag{3.1}$$

with x_1, m, y_1, y_2 satisfying (1.7) and (1.8), and for a given m with $M<m\leqq 2M$ let $r(n, m)$ denote the number of solutions of (3.1) with x_1, y_1, y_2 satisfying (1.7) and (1.8).

Let $s(n)=1$ or 0 according as $r(n)>0$ or $r(n)=0$. Then, by Cauchy's inequality

$$\sum_n r(n)=\sum_n s(n)\,r(n)\leqq\Big(\sum_n s(n)\Big)^{\frac{1}{2}}\Big(\sum_n r(n)^2\Big)^{\frac{1}{2}}.$$

Take $P=\frac{1}{3}N^{\frac{1}{3}}$. Then

$$\sum_n s(n)\leqq\mathscr{N}(N),\quad \sum_n r(n)\gg P^{\frac{20}{7}}$$

and by Cauchy's inequality

$$\sum_n r(n)^2=\sum_n\Big(\sum_m r(n, m)\Big)^2\ll P^{\frac{1}{7}}\sum_n\sum_m r(n, m)^2=P^{\frac{1}{7}}S.$$

Hence

$$P^{\frac{20}{7}}\ll\mathscr{N}(N)^{\frac{1}{2}}(P^{\frac{1}{7}}S)^{\frac{1}{2}}$$

and so Theorem 6 follows from Theorem A.

4. The proof of Theorem B

We wish to use Theorem A. However the presence of the factor P^{ε} is a nuisance when M is small. We therefore require a number of refinements of standard estimates such as Weyl's inequality and Hua's Lemma. These refinements are largely based on the work of Hooley [H1], [H2].

Lemma 1. *Let $\mathfrak{m}(R)$ denote the set of real numbers α with the property that whenever $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}R^{-\frac{9}{4}}$ and $(a, q)=1$ we have $q>R^{\frac{3}{4}}$, and define*

$$f(\alpha; R)=\sum_{x\leqq R} e(\alpha x^3).$$

Then, uniformly for $\alpha\in\mathfrak{m}(R)$ we have

$$f(\alpha; R)\ll R^{\frac{3}{4}}(\log R)^{\frac{1}{p}+\varepsilon}.$$

Proof. Let

$$\delta=10^{-2}.$$

Let $\alpha \in \mathfrak{m}(R)$ and choose a, q so that $(a, q)=1$, $q \leqq R^{2-\delta}$, $\left|\alpha-\frac{a}{q}\right| \leqq q^{-1} R^{\delta-2}$. If $q \leqq R^{\frac{3}{2}-\delta}$, then by Theorem 2 of Vaughan [V2] and Lemma 4.6 of Vaughan [V1]

$$f(\alpha; R) \ll q^{-\frac{1}{3}} R\left(1+R^3\left|\alpha-\frac{a}{q}\right|\right)^{-\frac{1}{3}}+q^{\frac{1}{2}+\varepsilon}\left(1+R^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{2}}$$

$$\ll q^{-\frac{1}{3}} R\left(1+R^3\left|\alpha-\frac{a}{q}\right|\right)^{-\frac{1}{3}}+R^{\frac{3}{4}}.$$

If $q>R^{\frac{3}{4}}$, or if $q \leqq R^{\frac{3}{4}}$ and $\left|\alpha-\frac{a}{q}\right|>q^{-1} R^{-\frac{9}{4}}$ this is

$$\ll R^{\frac{3}{4}}.$$

Since $\alpha \in \mathfrak{m}(R)$ this exhausts the possibilities when $q \leqq R^{\frac{3}{2}-\delta}$.

Now suppose that $q>R^{\frac{3}{2}-\delta}$. Then $\left|\alpha-\frac{a}{q}\right| \leqq R^{2\delta-\frac{7}{2}}$. Therefore

$$f(\alpha; R)=f\left(\frac{a}{q}; R\right)+O(R^{\frac{1}{2}+2\delta}).$$

We now refine the proof of Weyl's inequality. By the standard differencing technique applied via Cauchy's inequality we have

$$f\left(\frac{a}{q}; R\right)^4 \ll R^3+R \sum_{h \leqq R} \sum_{j \leqq 6R} \min\left(R, \left\|\frac{ahj}{q}\right\|^{-1}\right).$$

The double sum on the right is

$$\sum_{|l| \leqq \frac{1}{2}q} \min\left(R, \frac{q}{|l|}\right) \sum_{\substack{h \leqq R \\ ahj \equiv l \,(\mathrm{mod}\, q)}} \sum_{j \leqq 6R} 1.$$

By Theorem 3 of Hooley [H1] this is

$$\ll \sum_{|l| \leqq \frac{1}{2}q} \min\left(R, \frac{q}{|l|}\right) d((q, l))\, q^{-1} R^2 (\log R)^{\frac{4}{\pi}-1+\varepsilon}$$

$$\ll R^2 (\log R)^{\frac{4}{\pi}-1+\varepsilon}\left(Rd(q)q^{-1}+\sum_{r|q} \frac{d(r)}{r} \sum_{m \leqq \frac{q}{r}} \frac{1}{m}\right)$$

$$\ll R^2 (\log R)^{\frac{4}{\pi}-1+\varepsilon}\left(1+(\log\log q)^2 \log q\right).$$

This gives the desired conclusion.

Lemma 2. *Let $\mathscr{A}$ be a set of natural numbers not exceeding R with cardinality A. Then the number T of solutions of*

$$x_1^3+x_2^3=x_3^3+x_4^3$$

with $x_i \in \mathscr{A}$ satisfies

$$T \ll A^{2-\varepsilon} R^{\varepsilon}.$$

In particular

$$T \ll R^2.$$

Proof. Let $r(n)$ denote the number of representations of n as the sum of two positive cubes and let $r(n; \mathscr{A})$ denote the number of solutions of

$$x_1^3 + x_2^3 = n$$

with $x_i \in \mathscr{A}$. Then, by Hölder's inequality

$$T = \sum_n r(n; \mathscr{A})^2 \leqq \Big(\sum_n r(n; \mathscr{A})\Big)^{1-\frac{\varepsilon}{2}} \Big(\sum_{n \leqq 2R^3} r(n)^{\frac{2}{\varepsilon}+1}\Big)^{\frac{\varepsilon}{2}}.$$

Hence, by Theorem 2 of Hooley [H2] we have

$$T \ll (A^2)^{1-\frac{\varepsilon}{2}} (R^2)^{\frac{\varepsilon}{2}} = A^{2-\varepsilon} R^{\varepsilon}.$$

Lemma 3. *Let $\mathscr{A}$ be as in Lemma 2 and let U denote the number of solutions of*

$$x_1^3 + x_2^3 + y_1^3 + y_2^3 = x_3^3 + x_4^3 + y_3^3 + y_4^3$$

with $x_i \in \mathscr{A}$ and $y_i \leqq R$. Then

$$U \ll A^{2-\varepsilon} R^{3+\varepsilon} (\log R)^{\frac{4}{\pi}+\varepsilon}.$$

Proof. Let

$$f(\alpha) = \sum_{y \leqq R} e(\alpha y^3), \quad g(\alpha) = \sum_{x \in \mathscr{A}} e(\alpha x^3).$$

Then

$$U = \int_0^1 |f(\alpha)^4 g(\alpha)^4| \, d\alpha.$$

Let $\mathfrak{m} = \mathfrak{m}(R) \cap (0, 1]$ where $\mathfrak{m}(R)$ is as in Lemma 1, and let $\mathfrak{M} = (0, 1] \setminus \mathfrak{m}$. Then

$$U = U_{\mathfrak{M}} + U_{\mathfrak{m}}$$

where

$$U_{\mathscr{B}} = \int_{\mathscr{B}} |f(\alpha)^4 g(\alpha)^4| \, d\alpha.$$

Hence either $U \leqq 2U_{\mathfrak{m}}$ or $U \leqq 2U_{\mathfrak{M}}$. In the first case, by Lemmas 1 and 2 we have

$$U \ll TR^3 (\log R)^{\frac{4}{\pi}+\varepsilon} \ll A^{2-\varepsilon} R^{3+\varepsilon} (\log R)^{\frac{4}{\pi}+\varepsilon}.$$

Hence we may suppose that

$$U \leqq 2U_{\mathfrak{M}}. \tag{4.1}$$

By a standard application of the Hardy-Littlewood method we have

$$\int_{\mathfrak{M}} |f(\alpha)|^{6+\varepsilon} d\alpha \ll R^{3+\varepsilon}.$$

Therefore, by Hölder's inequality

$$U_{\mathfrak{M}} \leqq A^{\frac{8-4\varepsilon}{6+\varepsilon}} \left(\int_{\mathfrak{M}} |f(\alpha)|^{6+\varepsilon} d\alpha\right)^{\frac{4}{6+\varepsilon}} \left(\int_0^1 |g(\alpha)|^8 d\alpha\right)^{\frac{2+\varepsilon}{6+\varepsilon}}$$

$$\ll A^{\frac{8-4\varepsilon}{6+\varepsilon}} R^{\frac{12+4\varepsilon}{6+\varepsilon}} \left(\int_0^1 |g(\alpha)|^8 d\alpha\right)^{\frac{2+\varepsilon}{6+\varepsilon}}.$$

Obviously the integral on the right does not exceed U. Hence, by (4. 1),

$$U \ll A^{\frac{8-4\varepsilon}{6+\varepsilon}} R^{\frac{12+4\varepsilon}{6+\varepsilon}} U^{\frac{2+\varepsilon}{6+\varepsilon}}$$

whence

$$U \ll A^{2-\varepsilon} R^{3+\varepsilon}$$

which is more than sufficient.

The following lemma is a refinement of (2. 14).

Lemma 4. *Let* $\delta = 10^{-3}$, $1 \leqq H \leqq P$ *and let* $\mathfrak{n}$ *denote the set of* α *in* $(0, 1]$ *such that whenever* $\left|\alpha - \frac{a}{q}\right| \leqq q^{-1} H^{-1} P^{\delta-1}$ *and* $(a, q) = 1$ *we have* $q \geqq P^{1+\delta}$. *Then*

$$\sum_{h \leqq H} \left|\sum_{x \leqq P} e(3\alpha h x^2)\right|^2 \ll HP(\log P)^{\frac{4}{\pi}+\varepsilon}$$

uniformly for $\alpha \in \mathfrak{n}$.

Proof. We follow the pattern of the proof of Lemma 1. Let $\alpha \in \mathfrak{n}$ and choose a, q so that $\left|\alpha - \frac{a}{q}\right| \leqq q^{-1} H^{-1} P^{\delta-1}$, $(a, q) = 1$ and $q \leqq HP^{1-\delta}$. Then, by the definition of $\mathfrak{n}$, $q \geqq P^{1+\delta}$. Let

$$F(\gamma; X) = \sum_{x \leqq X} e(\gamma x^2).$$

Then, by Lemma 2. 6 of Vaughan [V1],

$$F(3\alpha h; P) = e(3\beta h P^2) F\left(3\frac{ah}{q}; P\right) - \int_0^P 12\pi i \beta h X e(3\beta X^2) F\left(3\frac{ah}{q}; X\right) dX$$

where

$$\beta = \alpha - \frac{a}{q}.$$

Therefore, by Cauchy's inequality

$$|F(3\alpha h; P)|^2 \ll \left|F\left(3\frac{ah}{q}; P\right)\right|^2 + P^{-1} \int_0^P \left|F\left(3\frac{ah}{q}; X\right)\right|^2 dX.$$

Hence, by (2. 14) (with P replaced by X),

$$\sum_{h\leqq H}\Big|\sum_{x\leqq P}e(3\alpha hx^2)\Big|^2$$
$$\ll HP^{1-\frac{1}{2}\delta+\varepsilon}+\sup_{P^{1-\frac{1}{2}\delta}\leqq X\leqq P}\left(\sum_{h\leqq H}\left|F\left(3\frac{ah}{q};X\right)\right|^2\right).$$

Thus it suffices to estimate

$$\sum_{h\leqq H}\left|F\left(3\frac{ah}{q};X\right)\right|^2$$

when $P^{1-\frac{1}{2}\delta}\leqq X\leqq P$. This is

$$\ll HX+\sum_{h\leqq H}\sum_{j\leqq 6X}\min\left(X,\left\|\frac{ahj}{q}\right\|^{-1}\right)$$
$$\ll HP+\sum_{h\leqq H}\sum_{j\leqq 6P}\min\left(P,\left\|\frac{ahj}{q}\right\|^{-1}\right).$$

As in the proof of Lemma 1, since $HP^{1-\delta}\leqq(HP)^{1-\frac{1}{2}\delta}$ this is

$$\ll HP(\log P)^{\frac{4}{\pi}+\varepsilon}$$

as required.

We are now in a position to establish the following refinement of Theorem A.

Lemma 5. *Suppose that $M\leqq P^{\frac{1}{5}}$. Then in the notation of Theorem A we have*

$$S\ll MPQ^2(\log P)^5\,(P^{\frac{1}{2}}M^{-\frac{7}{2}}+1). \tag{4. 2}$$

In particular, when $M\leqq P^{\frac{1}{7}}$ we have

$$S\ll P^{\frac{7}{2}}(\log P)^5M^{-\frac{9}{2}}. \tag{4. 3}$$

Proof. We simply indicate the amendments that have to be made to the proof of Theorem A.

When $(u, m)=1$ the number of solutions of

$$z^3\equiv u\,(\mathrm{mod}\,m^3)$$

is $\ll 3^{\omega(m)}$ where $\omega(m)$ denotes the number of different prime divisors of m. Therefore

$$S\ll\sum_{M<m\leqq 2M}3^{\omega(m)}S_0(m)$$

where $S_0(m)$ is the number of solutions of (1. 6), (1. 7), (1. 8) subject to (2. 3). By Lemma 2, for a given m the number of solutions with $x_1=x_2$ is $\ll PQ^2$ and we have

$$\sum_{M<m\leqq 2M}3^{\omega(m)}\ll M(\log P)^2. \tag{4. 4}$$

Thus it suffices to show that

$$S_1 \ll MPQ^2(\log P)^5\,(P^{\frac{1}{2}}M^{-\frac{7}{2}}+1)$$

where

$$S_1 = \sum_{M<m\leqq 2M} 3^{\omega(m)}S_1(m)$$

and $S_1(m)$ is the number of solutions of (1. 6), (1. 7), (1. 8), (2. 3) with $x_2 > x_1$. Proceeding as in § 2 we obtain

$$S_1 \leqq S_2$$

where

$$S_2 = \int_0^1 \sum_{h\leqq H} F(3\alpha h)\,G(\alpha h^3)\,|g(4\alpha)|^4\,d\alpha$$

with F and g as in (2. 9) and (2. 2) respectively, and

$$G(\beta) = \sum_{M<m\leqq 2M} 3^{\omega(m)}e(\beta m^6).$$

Let $\delta = 10^{-3}$, let

$$\mathfrak{M}(q,a) = \left\{\alpha: \left|\alpha - \frac{a}{q}\right| \leqq (qP^{1-\delta}H)^{-1}\right\},$$

let $\mathfrak{M}$ denote the union of the $\mathfrak{M}(q,a)$ with $1\leqq a\leqq q\leqq P^{1+\delta}$ and $(a,q)=1$, let $\mathfrak{U} = (P^{\delta-1}H^{-1}, 1+P^{\delta-1}H^{-1}]$, and let $\mathfrak{m} = \mathfrak{U}\setminus\mathfrak{M}$.

In the treatment of $I_{\mathfrak{M}}$ we observe that for $\alpha\in\mathfrak{M}(q,a)$ we have, by (2. 14),

$$\sum_{h\leqq H}|F(3\alpha h)|^2 \ll HP^{2+\varepsilon}q^{-1}.$$

Moreover

$$g(4\alpha) \ll q^{-\frac{1}{3}}Q\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{-\frac{1}{3}} + P^{\frac{1}{2}+\frac{1}{2}\delta+\varepsilon}.$$

Therefore

$$\begin{aligned} I_{\mathfrak{M}} &\ll \sum_{q\leqq P^{1+\delta}} HMP^{1+\delta}q^{\frac{1}{2}}(Qq^{-\frac{4}{3}} + P^{2+2\delta}q^{-1}HP^{\delta-1}) \\ &\ll MP^{1+5\delta}Q^2(P^{\frac{1}{6}}M^{-2} + M^2P^{-\frac{1}{2}}) \\ &\ll MPQ^2(P^{\frac{1}{2}}M^{-\frac{7}{2}}+1). \end{aligned}$$

It remains to treat $I_{\mathfrak{m}}$. In order to estimate I, given by (2. 17), we observe that the $\mathfrak{m}$ under consideration is precisely the $\mathfrak{n}$ of Lemma 4. Thus, by Lemma 2,

$$I \ll HP(\log P)^{\frac{4}{\pi}+\varepsilon}Q^2,$$

whence, cf. (2. 19),

$$I_{\mathfrak{m}} \ll H^{\frac{1}{2}}P^{\frac{1}{2}}(\log P)^{\frac{2}{\pi}+\varepsilon}QJ^{\frac{1}{2}}. \tag{4. 6}$$

We proceed to estimate J as in § 2. Now

$$\sum_{h\leqq H}|G(\alpha h^3)|^2 \ll H\sum_{m\leqq 2M} 9^{\omega(m)} + \sum\sum_{M<m_1<m_2\leqq 2M} 3^{\omega(m_1)+\omega(m_2)}\,|E(\alpha m_2^6-\alpha m_1^6)|.$$

Therefore

$$J \ll HM(\log P)^8 Q^2 + \sum\sum_{M<m_1<m_2\leqq 2M} 3^{\omega(m_1)+\omega(m_2)} K(m_2^6-m_1^6). \tag{4.7}$$

Let $\mathfrak{N}$ and $\mathfrak{n}$ be as in § 2. Then, in the notation of Lemma 1, $\mathfrak{n}\subset\mathfrak{m}(H)$. Thus for $\beta\in\mathfrak{n}$ we have

$$E(\beta)\ll H^{\frac{3}{4}}(\log H)^{\frac{1}{\pi}+\varepsilon},$$

so that

$$K(d)=d^{-1}\int_{\mathfrak{N}}|E(\beta)|\left|g\left(\frac{4\beta}{d}\right)\right|^4 d\beta + O\left(H^{\frac{3}{4}}(\log P)^{\frac{1}{\pi}+\varepsilon}Q^2\right). \tag{4.8}$$

This leads to the estimate

$$K(d)\ll H^{\frac{1}{4}}Q^{\frac{9}{4}}P^{\varepsilon} + H^{\frac{3}{4}}(\log P)^{\frac{1}{\pi}+\varepsilon}Q^2.$$

Hence, by (4.4), (4.7) and (4.8),

$$\begin{aligned} I_{\mathfrak{m}} &\ll H^{\frac{1}{2}}P^{\frac{1}{2}}(\log P)^{\frac{2}{\pi}+\varepsilon}Q\left(H^{\frac{1}{2}}M^{\frac{1}{2}}(\log P)^4 Q + MH^{\frac{1}{8}}Q^{\frac{9}{8}}P^{\varepsilon} + M(\log P)^{2+\frac{1}{2\pi}}H^{\frac{3}{8}}Q\right)\\ &\ll MPQ^2(\log P)^5\left(P^{\frac{1}{2}}M^{-\frac{7}{2}} + P^{\frac{1}{4}+\varepsilon}M^{-2} + P^{\frac{3}{8}}M^{-\frac{21}{8}}\right) \end{aligned}$$

and the lemma is proved.

We now proceed to complete the proof of Theorem 3. For brevity we write

$$L=\log P$$

and let $m(z)$ denote the smallest prime divisor p of z with $p>L^{80}$ if one exists, and $+\infty$ otherwise.

We define five sets of ordered pairs x, y as follows.

$$\begin{aligned} \mathscr{C}_d &= \{x, y:\ x\leqq P,\ y\leqq P,\ (x,y)=d\},\\ \mathscr{C} &= \bigcup_{d>L^{80}}\mathscr{C}_d,\\ \mathscr{D} &= \{x, y:\ x\leqq P,\ y\leqq P,\ (x,y)\leqq L^{80},\ m(y)\leqq P^{\frac{1}{7}}\},\\ \mathscr{E} &= \{x, y:\ x\leqq P,\ y\leqq P,\ (x,y)\leqq L^{80},\ m(y)>P^{\frac{1}{7}},\ m(x)\leqq P^{\frac{1}{7}}\},\\ \mathscr{J} &= \{x, y:\ x\leqq P,\ y\leqq P,\ (x,y)\leqq L^{80},\ m(y)>P^{\frac{1}{7}},\ m(x)>P^{\frac{1}{7}}\}. \end{aligned}$$

For $\mathscr{B}$ one of the above sets, let

$$I(\mathscr{B}) = \int_{\mathfrak{m}} \Big(\sum_{x,y \in \mathscr{B}} e(\alpha x^3 - \alpha y^3) \Big) |f(\alpha)|^6 \, d\alpha.$$

Then

$$\int_{\mathfrak{m}} |f(\alpha)|^8 \, d\alpha = I(\mathscr{C}) + I(\mathscr{D}) + I(\mathscr{E}) + I(\mathscr{F}).$$

By Hölder's inequality and two applications of Lemma 3 (the first with R replaced by $\frac{P}{d}$),

$$\begin{aligned} I(\mathscr{C}_d) &\leqq \left(\int_0^1 \Big| \sum_{x,y \in \mathscr{C}_d} e(\alpha x^3 - \alpha y^3) \Big|^4 d\alpha \right)^{\frac{1}{4}} \left(\int_0^1 |f(\alpha)|^8 \, d\alpha \right)^{\frac{3}{4}} \\ &\ll \left(\left(\frac{P}{d} \right)^5 L^{\frac{4}{\pi}+\varepsilon} \right)^{\frac{1}{4}} (P^5 L^{\frac{4}{\pi}+\varepsilon})^{\frac{3}{4}} \\ &= P^5 L^{\frac{4}{\pi}+\varepsilon} d^{-\frac{5}{4}}. \end{aligned}$$

Therefore

$$I(\mathscr{C}) = \sum_{d > L^{80}} I(\mathscr{C}_d) \ll P^5 (\log P)^{-2}.$$

To estimate $I(\mathscr{D})$ and $I(\mathscr{E})$ we argue as follows. Let $\mathscr{V} = \mathscr{D}$ or $\mathscr{E}$. Then, by Lemma 1,

$$I(\mathscr{V}) \ll P^{\frac{3}{4}} L^{\frac{1}{\pi}+\varepsilon} \int_0^1 \Big| \sum_{x,y \in \mathscr{V}} e(\alpha x^3 - \alpha y^3) \Big| \, |f(\alpha)|^5 \, d\alpha.$$

By Schwarz's inequality the integral on the right is bounded by

$$J(\mathscr{V})^{\frac{1}{2}} \left(\int_0^1 |f(\alpha)|^8 \, d\alpha \right)^{\frac{1}{2}}$$

where

$$J(\mathscr{V}) = \int_0^1 \Big| \sum_{x,y \in \mathscr{V}} e(\alpha x^3 - \alpha y^3) \Big|^2 \, |f(\alpha)|^2 \, d\alpha.$$

Clearly

$$J(\mathscr{V}) \leqq V$$

where V is the number of solutions of

$$x_1^3 - m_1^3 y_1^3 + z_1^3 = x_2^3 - m_2^3 y_2^3 + z_2^3$$

with $x_i \leqq P$, $y_i \leqq \frac{P}{m_i}$, $L^{80} < m_i \leqq P^{\frac{1}{7}}$, $(x_i, m_i) = 1$ (note that if $(x, y) \leqq L^{80}$, p divides y and $p > L^{80}$, then $p \nmid x$). Hence, by Lemma 3,

$$I(\mathscr{V}) \ll V^{\frac{1}{2}} P^{\frac{13}{4}} L^{\frac{3}{\pi}+\varepsilon}. \tag{4.9}$$

Let

$$\mathscr{M} = \{M:\ M = 2^k L^{80},\ k = 0, 1, \ldots;\ M \leqq P^{\frac{1}{7}}\}.$$

Then

$$V \leqq \int_0^1 \left| \sum_{M \in \mathscr{M}} \sum_{M < m \leqq 2M} f_m(\alpha) f\left(-\alpha m^3; \frac{P}{M}\right)\right|^2 |f(\alpha)|^2 d\alpha$$

where

$$f_m(\alpha) = \sum_{\substack{x \leqq P \\ (x,m)=1}} e(\alpha x^3), \quad f(\beta; Q) = \sum_{x \leqq Q} e(\beta x^3).$$

Therefore, by Cauchy's inequality and Hölder's inequality

$$V \ll L \sum_{M \in \mathscr{M}} M \int_0^1 \sum_{M < m \leqq 2M} |f_m(\alpha)|^2 \left| f\left(\alpha m^3; \frac{P}{M}\right)\right|^2 |f(\alpha)|^2 d\alpha$$

$$\ll L \sum_{M \in \mathscr{M}} M V_1(M)^{\frac{1}{2}} V_2(M)^{\frac{1}{4}} V_3(M)^{\frac{1}{4}}$$

where

$$V_1(M) = \int_0^1 \sum_{M < m \leqq 2M} |f_m(\alpha)|^2 \left| f\left(\alpha m^3; \frac{P}{M}\right)\right|^4 d\alpha,$$

$$V_2(M) = \int_0^1 \sum_{M < m \leqq 2M} |f_m(\alpha)|^4 d\alpha,$$

$$V_3(M) = \int_0^1 M |f(\alpha)|^8 d\alpha.$$

Hence, by Lemmas 5, 2 and 3,

$$V \ll L \sum_{M \in \mathscr{M}} M (P^{\frac{7}{2}} L^5 M^{-\frac{9}{2}})^{\frac{1}{2}} (MP^2)^{\frac{1}{4}} (MP^5 L^{\frac{4}{\pi}+\varepsilon})^{\frac{1}{4}}$$

$$\ll P^{\frac{7}{2}} L^4 \sum_{M \in \mathscr{M}} M^{-\frac{3}{4}} \ll P^{\frac{7}{2}} L^{-56}.$$

Therefore, by (4. 9),

$$I(\mathscr{V}) \ll P^5 L^{-2}.$$

It remains to treat $I(\mathscr{J})$. We now perform the above operations on the next two cubes. Thus

$$I(\mathscr{J}) = I(\mathscr{J}, \mathscr{C}) + I(\mathscr{J}, \mathscr{D}) + I(\mathscr{J}, \mathscr{E}) + I(\mathscr{J}, \mathscr{J})$$

where

$$I(\mathscr{J}, \mathscr{B}) = \int_{\mathfrak{m}} \Big(\sum_{x,y \in \mathscr{J}} e(\alpha x^3 - \alpha y^3)\Big) \Big(\sum_{x,y \in \mathscr{B}} e(\alpha x^3 - \alpha y^3)\Big) |f(\alpha)|^4 d\alpha.$$

By Hölder's inequality

$$I(\mathscr{J}, \mathscr{C}_d) \leqq \left(\int_0^1 |\sum_{x,y\in\mathscr{J}} e(\alpha x^3 - \alpha y^3)|^4 d\alpha\right)^{\frac{1}{4}} \left(\int_0^1 |\sum_{x,y\in\mathscr{C}_d} e(\alpha x^3 - \alpha y^3)|^4 d\alpha\right)^{\frac{1}{4}} \left(\int_0^1 |f(\alpha)|^8 d\alpha\right)^{\frac{1}{2}}$$

and so $I(\mathscr{J}, \mathscr{C})$ may be estimated in the same way as $I(\mathscr{C})$. For $\mathscr{V} = \mathscr{D}$ or $\mathscr{E}$ we have, by Lemma 1,

$$I(\mathscr{J}, \mathscr{V}) \ll P^{\frac{3}{4}} L^{\frac{1}{\pi}+\varepsilon} \int_0^1 |\sum_{x,y\in\mathscr{J}} e(\alpha x^3 - \alpha y^3)| \, |\sum_{x,y\in\mathscr{V}} e(\alpha x^3 - \alpha y^3)| \, |f(\alpha)|^3 d\alpha$$

and by Hölder's inequality this is

$$\ll P^{\frac{3}{4}} L^{\frac{1}{\pi}+\varepsilon} \left(\int_0^1 |\sum_{x,y\in\mathscr{J}} e(\alpha x^3 - \alpha y^3)|^4 d\alpha\right)^{\frac{1}{4}} J(\mathscr{V})^{\frac{1}{2}} \left(\int_0^1 |f(\alpha)|^8 d\alpha\right)^{\frac{1}{4}}$$

and so again we may proceed as for $I(\mathscr{D})$ and $I(\mathscr{E})$.

It remains to consider $I(\mathscr{J}, \mathscr{J})$. Clearly

$$|I(\mathscr{J}, \mathscr{J})| \leqq \int_0^1 |\sum_{x,y\in\mathscr{J}} e(\alpha x^3 - \alpha y^3)|^2 \, |f(\alpha)|^4 d\alpha$$

so that

$$|I(\mathscr{J}, \mathscr{J})| \leqq J$$

where J is the number of solutions of

$$x_1^3 + x_2^3 + y_1^3 + y_2^3 = x_3^3 + x_4^3 + y_3^3 + y_4^3 \tag{4.10}$$

with $y_i \leqq P$, $(x_1, x_3) \in \mathscr{J}$, $(x_2, x_4) \in \mathscr{J}$. By the definition of $\mathscr{J}$ no x_i can have a prime divisor p with $L^{80} < p \leqq P^{\frac{1}{7}}$. Hence $I(\mathscr{J}, \mathscr{J})$ is bounded by the number of solutions of (4.10) with $x_i \in \mathscr{A}$, $y_i \leqq P$ where $\mathscr{A}$ is the set of natural numbers not exceeding P with no prime divisor p in the range $L^{80} < p \leqq P^{\frac{1}{7}}$. Therefore, by Lemma 3,

$$I(\mathscr{J}, \mathscr{J}) \ll A^{2-\varepsilon} P^{3+\varepsilon} (\log P)^{\frac{4}{\pi}+\varepsilon}$$

where $A = \operatorname{card} \mathscr{A}$. By Theorem 3.3 of Halberstam and Richert [Ha],

$$A \ll \frac{P \log\log P}{\log P}.$$

Therefore

$$I(\mathscr{J}, \mathscr{J}) \ll P^5 (\log P)^{\frac{4}{\pi}-2+\varepsilon}.$$

This completes the proof of Theorem B.

5. A technical lemma

Suppose that C is a sufficiently large constant, that $\varepsilon > 0$ and that

$$P > P_0(\varepsilon), \quad P \in \mathbb{N}. \tag{5.1}$$

Let

$$Y = P^{\frac{17}{115}}, \quad H = CPY^{-3}, \quad Q = \frac{P}{Y}, \quad Z = Q^{\frac{1}{7}}, \quad R = \frac{Q}{Z}, \quad M = P^{\frac{27}{115}}, \tag{5.2}$$

and for primes p and p_1 with $Y < p \leqq 2Y$, $Z < p_1 \leqq 2Z$ define

$$f_p(\alpha) = \sum_{\substack{P < x \leqq 2P \\ p \nmid x}} e(\alpha x^3), \tag{5.3}$$

$$g_d(\alpha) = \sum_{\substack{Q < y \leqq 2Q \\ (d,y)=1}} e(\alpha y^3), \tag{5.4}$$

$$h(\alpha) = \sum_{R < z \leqq 2R} e(\alpha z^3), \tag{5.5}$$

$$S(\alpha) = \Big| \sum_{\substack{Z < p_1 \leqq 2Z \\ p_1 \equiv 2 \,(\mathrm{mod}\, 3)}} \sum_{\substack{Z < p_2 \leqq 2Z \\ p_2 \equiv 2 \,(\mathrm{mod}\, 3)}} g_{p_1 p_2}(\alpha)\, h(p_1^3 \alpha)\, h(p_2^3 \alpha) \Big|^2, \tag{5.6}$$

$$\Phi_p(\alpha) = \sum_{\substack{P < y \leqq 2P \\ p \nmid y}} 1 + 2\,\mathrm{Re} \sum_{h \leqq H} \sum_{\substack{2P + hp^3 < y \leqq 4P - hp^3 \\ p \nmid y,\ y \equiv h \,(\mathrm{mod}\, 2)}} e\left(\frac{3}{4}\alpha h y^2 + \frac{1}{4}\alpha h^3 p^6\right). \tag{5.7}$$

We now state a technical lemma which we apply in § 8 to establish a suitable minor arcs estimate for use in the proofs of Theorems 4 and 5.

Lemma 6. *Let* $\mathfrak{n}$ *denote the set of real numbers* α *in* $(0, 1]$ *with the property that whenever* $\left|\alpha - \frac{a}{q}\right| \leqq q^{-1} M Q^{-3}$ *and* $(a, q) = 1$ *we have* $q > M$, *and let*

$$T = \int_{\mathfrak{n}} \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \,(\mathrm{mod}\, 3)}} \Phi_p(\alpha)\, S(\alpha)\, d\alpha.$$

Then

$$T \ll YZ^4 P^{1+\varepsilon} Q R^2.$$

The first step in the proof of this lemma is to manipulate $\Phi_p(\alpha)$ so as to make it more amenable to treatment by our method. By (5.7),

$$\Phi_p(\alpha) = \Psi_p(\alpha) - \Xi_p(\alpha) + O(P) \tag{5.8}$$

where

$$\Psi_p(\alpha) = 2\,\mathrm{Re} \sum_{h \leqq H} F_p(\alpha; h)\, e\left(\frac{1}{4}\alpha h^3 p^6\right), \tag{5.9}$$

$$F_p(\alpha; h) = \sum_{\substack{2P+hp^3 < y \leqq 4P-hp^3 \\ y \equiv h \pmod 2}} e\left(\frac{3}{4}\alpha h y^2\right), \tag{5.10}$$

$$\Xi_p(\alpha) = 2\,\mathrm{Re} \sum_{h \leqq H} D_p(\alpha; h)\, e\left(\frac{1}{4}\alpha h^3 p^6\right), \tag{5.11}$$

$$D_p(\alpha; h) = \sum_{\substack{2\frac{P}{p}+hp^2 < y \leqq 4\frac{P}{p}-hp^2 \\ y \equiv h \pmod 2}} e\left(\frac{3}{4}\alpha h p^2 y^2\right). \tag{5.12}$$

Then

$$T = \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} \big(T_1(p) - T_2(p)\big) + O(YPT_3) \tag{5.13}$$

where

$$T_1(p) = \int_{\mathfrak{n}} \Psi_p(\alpha)\, S(\alpha)\, d\alpha, \tag{5.14}$$

$$T_2(p) = \int_{\mathfrak{n}} \Xi_p(\alpha)\, S(\alpha)\, d\alpha, \tag{5.15}$$

and

$$T_3 = \int_0^1 S(\alpha)\, d\alpha. \tag{5.16}$$

It follows from (5. 6), Cauchy's inequality, Schwarz's inequality and Theorem A, that

$$T_3 \ll Z^4 Q^{1+\varepsilon} R^2. \tag{5.17}$$

Therefore

$$T = \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} \big(T_1(p) - T_2(p)\big) + O(YZ^4 P^{1+\varepsilon} Q R^2).$$

The appearance of p in the summation conditions in (5. 10) is a nuisance, so we proceed to remove it by a technique introduced in § 5. 2 of Vaughan [V1]. Suppose that $hp^3 \leqq P$. Then, by (5. 10),

$$F_p(\alpha; h) = \int_0^1 K(\gamma, hp^3)\, F(\alpha h, \gamma; h)\, d\gamma$$

where

$$K(\gamma, r) = \sum_{2P+r < z \leqq 4P-r} e(\gamma z)$$

and

$$F(\beta, \gamma; h) = \sum_{\substack{2P < y \leqq 4P \\ y \equiv h \pmod 2}} e\left(\frac{3}{4}\beta y^2 - \gamma y\right).$$

Let

$$\eta = P^{-2}.$$

Then

$$F_p(\alpha; h) = \int_{\eta}^{1-\eta} K(\gamma, hp^3)\, F(\alpha h, \gamma; h)\, d\gamma + O(1). \tag{5.18}$$

Also, since $P \in \mathbb{Z}$, we have

$$K(\gamma, r) = \begin{cases} \dfrac{e(\gamma)}{e(\gamma)-1}\left(e\left(\gamma(4P-r)\right) - e\left(\gamma(2P+r)\right)\right) & (r \leqq P), \\ 0 & (r > P). \end{cases} \tag{5.19}$$

By (5.9) and (5.18),

$$\Psi_p(\alpha) = \Theta_p(\alpha) + O(H) \tag{5.20}$$

where

$$\Theta_p(\alpha) = 2\,\mathrm{Re} \int_{\eta}^{1-\eta} \sum_{h \leqq H} K(\gamma, hp^3)\, F(\alpha h, \gamma; h)\, e\left(\frac{1}{4}\alpha h^3 p^6\right) d\gamma. \tag{5.21}$$

Therefore, by (5.14), (5.16) and (5.17)

$$T_1(p) = T_4(p) + O(Z^4 P^{1+\varepsilon} Q R^2) \tag{5.22}$$

where

$$T_4(p) = \int_{\mathfrak{n}} \Theta_p(\alpha)\, S(\alpha)\, d\alpha. \tag{5.23}$$

By (5.21)

$$\sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} \Theta_p(\alpha) = 2\,\mathrm{Re} \int_{\eta}^{1-\eta} \sum_{h \leqq H} F(\alpha h, \gamma; h) \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} K(\gamma, hp^3)\, e\left(\frac{1}{4}\alpha h^3 p^6\right) d\gamma, \tag{5.24}$$

and, by (5.19),

$$\sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} K(\gamma, hp^3)\, e\left(\frac{1}{4}\alpha h^3 p^6\right) = \frac{e(\gamma)}{e(\gamma)-1}\left(e(4\gamma P)\, G_h(\alpha h^3, -\gamma h) - e(2\gamma P)\, G_h(\alpha h^3, \gamma h)\right)$$

where

$$G_h(p, \sigma) = \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3 \\ p \leqq (\frac{P}{2h})^{\frac{1}{3}}}} e\left(\frac{1}{4}\rho p^6 + \sigma p^3\right). \tag{5.25}$$

Therefore, by (5. 22), (5. 23) and (5. 24),

$$\sum_{\substack{Y<p\leqq 2Y\\ p\equiv 2 \pmod 3}} T_1(p) \ll (\log P) \sup_{\substack{0\leqq\gamma\leqq 1\\ \theta=\pm 1}} T_5(\gamma,\theta) + YZ^2P^{1+\varepsilon}QR^2 \tag{5.26}$$

where

$$T_5(\gamma,\theta)=\int_{\mathfrak{n}} \sum_{h\leqq H} |F(\alpha h,\gamma;h)\, G_h(\alpha h^3,\theta\gamma h)|\, S(\alpha)\,d\alpha. \tag{5.27}$$

In § 6 we show that

$$T_5(\gamma,\theta) \ll YZ^4P^{1+\varepsilon}QR^2 \tag{5.28}$$

uniformly for $\gamma\in[0,1]$ and $\theta=\pm 1$, and in § 7 we show that

$$\sum_{\substack{Y<p\leqq 2Y\\ p\equiv 2 \pmod 3}} T_2(p) \ll YZ^4P^{1+\varepsilon}QR^2. \tag{5.29}$$

This establishes Lemma 6.

6. The estimation of T_5

We now attempt to imitate the proof of (2. 5). However, their is no obvious analogue of (2. 23) which is as effective, and in view of this we have to accept a slightly inflated choice of Y, namely $P^{\frac{17}{115}}$, rather than $P^{\frac{1}{7}}$. It is possible that, through the use of ideas of Hooley [H3] and in particular the use of Deligne's estimates on the major arcs, the value $P^{\frac{1}{7}}$ for Y could be achieved.

We require estimates for F, G, g, h which we derive in the following lemmas.

Lemma 7. *Suppose that* $\left|\alpha-\frac{a}{q}\right|\leqq q^{-2}$ *and* $(a,q)=1$. *Then*

$$\sum_{h\leqq H} |F(\alpha h,\gamma;h)|^2 \ll P^{\varepsilon}\left(\frac{q^{-1}HP^2}{1+Q^3\left|\alpha-\frac{a}{q}\right|}+HP+q\right).$$

Proof. When $\left|\alpha-\frac{a}{q}\right|\leqq(12HP^2)^{-1}$ or $q>P$ the lemma follows in the same manner as the Lemma of Vaughan [V3]. Thus we may suppose that

$$\left|\alpha-\frac{a}{q}\right|>(12HP^2)^{-1} \quad \text{and} \quad q\leqq P.$$

As in the proof of the Lemma of Vaughan [V3] we have

$$\sum_{h\leqq H}|F(\alpha h,\gamma;h)|^2\ll HP+P^{\varepsilon}\sum_{j\leqq 6HP}\min(P,\|\alpha j\|^{-1})$$

where $\|\beta\|$ is the distance of β from a nearest integer. When $j\leqq 6H$ and $q\nmid j$ we have

$$\|\alpha j\|=\left\|\frac{aj}{q}+j\left(\alpha-\frac{a}{q}\right)\right\|\geqq\left\|\frac{aj}{q}\right\|-6HP\left|\alpha-\frac{a}{q}\right|$$

$$\geqq\left\|\frac{aj}{q}\right\|-\frac{1}{2q}\geqq\frac{1}{2}\left\|\frac{aj}{q}\right\|.$$

Moreover, when $q|j$ we have $\|\alpha j\|=\left\|j\left(\alpha-\frac{a}{q}\right)\right\|=\left|j\left(\alpha-\frac{a}{q}\right)\right|$. Therefore

$$\sum_{j\leqq 6HP}\min(P,\|\alpha j\|^{-1})$$

$$\ll\sum_{r=1}^{q}(HPq^{-1}+1)\left\|\frac{ar}{q}\right\|^{-1}+\sum_{k\leqq\frac{6HP}{q}}\left(kq\left|\alpha-\frac{a}{q}\right|\right)^{-1}$$

$$\ll P^{\varepsilon}\left(HP+q+q^{-1}\left|\alpha-\frac{a}{q}\right|^{-1}\right).$$

We conclude the proof by observing that $HP^2=Q^3$.

Lemma 8. *Suppose that* $\alpha\in\mathbb{R}$, $\gamma\in\mathbb{R}$, *and* a, q *are such that* $(a,q)=1$, $q\leqq Q^3H^{-\frac{3}{4}}$, $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}H^{\frac{3}{4}}Q^{-3}$. *Then*

$$\sum_{h\leqq H}|G_h(\alpha h^3,\gamma h)|^2\ll P^{\varepsilon}\left(\frac{HY^2q^{-\frac{1}{3}}}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{3}}}+H^{\frac{3}{4}}Y^2\right).$$

Proof. By (5.25),

$$|G_h(\rho,\sigma)|^2=\sum_{\substack{Y<p_1\leqq 2Y\\ p_1\equiv 2\,(\mathrm{mod}\,3)\\ p_1^3\leqq\frac{P}{2h}}}\sum_{\substack{Y<p_2\leqq 2Y\\ p_2\equiv 2\,(\mathrm{mod}\,3)\\ p_2^3\leqq\frac{P}{2h}}}e\left(\frac{1}{4}\rho(p_2^6-p_1^6)+\sigma(p_2^3-p_1^3)\right).$$

Therefore,

$$\sum_{h\leqq H}|G_h(\alpha h^3,\gamma h)|^2\ll HY+\sum_{Y<p_1<p_2\leqq 2Y}\sum\left|E_{p_2}\left(\frac{1}{4}\alpha(p_2^6-p_1^6),\gamma(p_1^3-p_2^3)\right)\right| \tag{6.1}$$

where

$$E_p(\rho,\sigma)=\sum_{\substack{h\leqq H\\ h\leqq\frac{P}{2p^3}}}e(\rho h^3+\sigma h). \tag{6.2}$$

Since

$$YH = Y^2 H^{\frac{3}{4}} (P^{\frac{1}{4}} Y^{-\frac{7}{4}}) \leqq Y^2 H^{\frac{3}{4}}$$

we may concentrate our attention on the double sum in (6. 1). For brevity write

$$d = d(p_2, p_1) = p_2^6 - p_1^6 \tag{6. 3}$$

when $Y < p_1 < p_2 \leqq 2Y$. For such a d choose $c = c(d)$, $s = s(d)$ so that

$$\left| \frac{1}{4} \alpha d - \frac{c}{s} \right| \leqq s^{-1} (12 H^2)^{-1}, \quad (c, s) = 1, \quad s \leqq 12 H^2.$$

If $s(d) > H$, then Weyl's inequality gives

$$E_p \left(\frac{1}{4} \alpha d, \sigma \right) \ll H^{\frac{3}{4} + \varepsilon}$$

and so the total contribution from those p_2, p_1 for which $s(d) > H$ is

$$\ll P^{\varepsilon} Y^2 H^{\frac{3}{4}}.$$

Thus we may suppose that $s(d) \leqq H$. Let $I = \dfrac{P}{2p_2^3}$. Note that $I < H$. Also let $\beta = \dfrac{1}{4} \alpha d - \dfrac{c}{s}$. Then

$$E_{p_2} \left(\frac{1}{4} \alpha d, \sigma \right) = \sum_{h \leqq I} e\left(\left(\frac{c}{s} + \beta \right) h^3 + \sigma h \right)$$

$$= \frac{1}{s} \sum_{t=1}^{s} \left(\sum_{r=1}^{s} e\left(\frac{c}{s} r^3 + \frac{t}{s} r \right) \right) \sum_{h \leqq I} e\left(\beta h^3 + \left(\sigma - \frac{t}{s} \right) h \right).$$

By Theorem 7. 1 of Vaughan [V 1],

$$\sum_{r=1}^{s} e\left(\frac{c}{s} r^3 + \frac{t}{s} r \right) \ll s^{\frac{2}{3} + \varepsilon}.$$

Hence

$$E_{p_2} \left(\frac{1}{4} \alpha d, \sigma \right) \ll s^{\varepsilon - \frac{1}{3}} \sum_{s(\sigma - \frac{1}{2}) < t \leqq s(\sigma + \frac{1}{2})} \left| \sum_{h \leqq I} e\left(\beta h^3 + \left(\sigma - \frac{t}{s} \right) h \right) \right|.$$

For h and t in the expression on the right we have

$$\left| 3\beta h^2 + \sigma - \frac{t}{s} \right| \leqq \frac{3}{4}.$$

Moreover if

$$\left| \sigma - \frac{t}{s} \right| \geqq \frac{1}{2s},$$

then

$$\left|3\beta h^2+\sigma-\frac{t}{s}\right| \geqq \left|\sigma-\frac{t}{s}\right|-\frac{1}{4s} \geqq \frac{1}{2}\left|\sigma-\frac{t}{s}\right|,$$

so that by Lemmas 4. 2 and 4. 8 of Titchmarsh [T] we have

$$\sum_{h \leqq I} e\left(\beta h^3+\left(\sigma-\frac{t}{s}\right)h\right) \ll \left|\sigma-\frac{t}{s}\right|^{-1}.$$

Therefore

$$E_{p_2}\left(\frac{1}{4}\alpha d, \sigma\right) \ll s^{\varepsilon-\frac{1}{3}}\left|\sum_{h \leqq I} e\left(\beta h^3+\left(\sigma-\frac{t_0}{s}\right)h\right)\right|+H^{\frac{2}{3}+\varepsilon}$$

where the first term on the right only occurs when there is a t_0 with

$$\left|\sigma-\frac{t_0}{s}\right|<\frac{1}{2s}.$$

By Lemma 4. 2 of Titchmarsh [T],

$$\sum_{h \leqq I} e\left(\beta h^3+\left(\sigma-\frac{t_0}{s}\right)h\right)=\int_0^I e\left(\beta x^3+\left(\sigma-\frac{t_0}{s}\right)x\right)dx+O(1)$$

and by Theorem 7. 3 of Vaughan [V1]

$$\int_0^I e\left(\beta x^3+\left(\sigma-\frac{t_0}{s}\right)x\right)dx \ll |\beta|^{-\frac{1}{3}}.$$

Therefore, by (6. 4),

$$E_{p_2}\left(\frac{1}{4}\alpha d, \gamma(p_2^3-p_1^3)\right) \ll s^{\varepsilon-\frac{1}{3}} \min\left(H, \left|\frac{1}{4}\alpha d-\frac{c}{s}\right|^{-\frac{1}{3}}\right)+H^{\frac{3}{4}}.$$

The second term here is acceptable. Moreover if $s>H^{\frac{3}{4}}$, or $s \leqq H^{\frac{3}{4}}$ and

$$s\left|\frac{1}{4}\alpha d-\frac{c}{s}\right|>H^{-\frac{9}{4}},$$

then the first term is $\ll H^{\frac{3}{4}+\varepsilon}$, which is also acceptable. Thus it remains to estimate

$$\sum_{\substack{Y<p_1<p_2 \leqq 2Y \\ d=p_2^6-p_1^6 \in \mathscr{D}}}\sum s^{\varepsilon-\frac{1}{3}} \min\left(H, \left|\alpha d-\frac{4c}{s}\right|^{-\frac{1}{3}}\right)$$

where $\mathscr{D}$ is the set of numbers d such that there exist $s=s(d)$, $c=c(d)$ with $(s, c)=1$, $s \leqq H^{\frac{3}{4}}$, $\left|\frac{1}{4}\alpha d-\frac{c}{s}\right| \leqq s^{-1}H^{-\frac{9}{4}}$. The above sum is bounded by

$$\sum_{h \leqq Y}\sum_{p \in \mathscr{P}_h} s^{\varepsilon-\frac{1}{3}} \min\left(H, \left|\alpha(p+h)^6-\alpha p^6-\frac{4c}{s}\right|^{-\frac{1}{3}}\right)$$

where $\mathscr{P}_h$ is the set of primes p for which $p+h$ is prime, $Y<p \leqq 2Y$, $Y<p+h \leqq 2Y$, $(p+h)^6-p^6 \in \mathscr{D}$ and, of course, $s=s\big((p+h)^6-p^6\big)$, $a=a\big((p+h)^6-p^6\big)$.

For a given $h \leqq Y$ choose $r = r(h)$, $b = b(h)$ so that

$$\left|\alpha h - \frac{b}{r}\right| \leqq 8r^{-1}H^{-\frac{9}{4}}, \quad (b, r) = 1, \quad r \leqq \frac{1}{8}H^{\frac{9}{4}},$$

and for a given $p \in \mathscr{P}_h$ write $m = m(p) = h^{-1}((p+h)^6 - p^6)$. Then $6Y^5 \leqq m \leqq 6(2Y)^5$. Therefore

$$\left|\alpha h - \frac{4c}{sm}\right| \leqq 4(smH^{\frac{9}{4}})^{-1}$$

and

$$sm \leqq 6(2Y)^5 H^{\frac{3}{4}}.$$

Hence

$$|bsm - 4cr| \leqq 4rH^{-\frac{9}{4}} + 48(2Y)^5 H^{-\frac{3}{2}} < 1.$$

Therefore $\dfrac{4c}{sm} = \dfrac{b}{r}$ and so the sum in question is

$$\ll P^{\varepsilon} \sum_{h \leqq Y} r^{-\frac{1}{3}} \min\left(H, Y^{-\frac{5}{3}}\left|\alpha h - \frac{b}{r}\right|^{-\frac{1}{3}}\right) \sum_{p \in \mathscr{P}_h} \left(\frac{r}{s}\right)^{\frac{1}{3}}. \tag{6.5}$$

Since $(s, c) = 1$ and $(b, r) = 1$ we have $r = tu$ with $t \mid m$, $u \mid s$. Thus the innermost sum above is

$$\ll \sum_{t \mid r} t^{\frac{1}{3}} \sum_{\substack{p \in \mathscr{P}_h \\ t \mid m(p)}} 1. \tag{6.6}$$

The number of terms in the innermost sum here with $p \mid t$ or $p + h \mid t$ is $\ll P^{\varepsilon}$. Thus

$$\sum_{\substack{p \in \mathscr{P}_h \\ t \mid m(p)}} 1 \ll P^{\varepsilon} + \left(\frac{Y}{t} + 1\right) \sum_{\substack{x=1 \\ (x,t) = (x+h,t) = 1 \\ h^{-1}((x+h)^6 - x^6) \equiv 0 \,(\mathrm{mod}\, t)}}^{t} 1.$$

It is easily shown that the sum on the right here is $\ll t^{\varepsilon}$. Therefore

$$\sum_{\substack{p \in \mathscr{P}_h \\ t \mid m(p)}} 1 \ll P^{\varepsilon}\left(\frac{Y}{t} + 1\right).$$

Hence, by (6.5) and (6.6), the sum in question is

$$\ll P^{\varepsilon} \sum_{h \leqq Y} r^{-\frac{1}{3}} \min\left(H, Y^{-\frac{5}{3}}\left|\alpha h - \frac{b}{r}\right|^{-\frac{1}{3}}\right)(Y + r^{\frac{1}{3}})$$

$$\ll P^{\varepsilon} \sum_{h \leqq Y} r^{-\frac{1}{3}} \min\left(HY, Y^{-\frac{2}{3}}\left|\alpha h - \frac{b}{r}\right|^{-\frac{1}{3}}\right) + P^{\varepsilon} YH.$$

The term $P^{\varepsilon}YH$ is again acceptable and so can be ignored. The contribution from the h for which $r=r(h)>H^{\frac{3}{4}}$, or for which $r\leqq H^{\frac{3}{4}}$ and $\left|\alpha h-\frac{b}{r}\right|^{-1}<rH^{\frac{9}{4}}Y^5$ is also acceptable. Therefore it now remains to consider

$$\sum_{\substack{h\leqq Y\\ h\in\mathscr{H}}} r^{-\frac{1}{3}}\min\left(HY,\ Y^{-\frac{2}{3}}\left|\alpha h-\frac{b}{r}\right|^{-\frac{1}{3}}\right)$$

where $\mathscr{H}$ is the set of h for which there exist $r=r(h)$, $b=b(h)$ with $(r,b)=1$, $r\leqq H^{\frac{3}{4}}$,

$$\left|\alpha h-\frac{b}{r}\right|\leqq r^{-1}H^{-\frac{9}{4}}Y^{-5}.$$

Choose s, c so that $(s,c)=1$, $s\leqq\frac{1}{2}H^{\frac{9}{4}}Y^5$, $\left|\alpha-\frac{c}{s}\right|\leqq 2s^{-1}H^{-\frac{9}{4}}Y^{-5}$. Then for $h\in\mathscr{H}$ we have

$$|b(h)s-cr(h)h|\leqq\frac{1}{2}+2r(h)hH^{-\frac{9}{4}}Y^{-5}<1.$$

Therefore $\frac{b}{rh}=\frac{c}{s}$ and $s=rt$ with $t|h$, whence the above sum is

$$\begin{aligned}
&s^{-\frac{1}{3}}\sum_{\substack{h\leqq Y\\ h\in\mathscr{H}}}\left(\frac{s}{r}\right)^{\frac{1}{3}}\min\left(HY,\ Y^{-\frac{2}{3}}h^{-\frac{1}{3}}\left|\alpha-\frac{c}{s}\right|^{-\frac{1}{3}}\right)\\
&\qquad\ll s^{-\frac{1}{3}}\sum_{t|s}t^{\frac{1}{3}}\sum_{j\leqq\frac{Y}{t}}\min\left(HY,\ Y^{-\frac{2}{3}}j^{-\frac{1}{3}}t^{-\frac{1}{3}}\left|\alpha-\frac{c}{s}\right|^{-\frac{1}{3}}\right)\\
&\qquad\ll P^{\varepsilon}s^{-\frac{1}{3}}\min\left(HY^2,\ \left|\alpha-\frac{c}{s}\right|^{-\frac{1}{3}}\right).
\end{aligned}$$

If $s>H^{\frac{3}{4}}$ or $s\left|\alpha-\frac{c}{s}\right|>H^{-\frac{9}{4}}Y^{-6}$, then once more we obtain an acceptable estimate. Hence we may suppose that $s\leqq H^{\frac{3}{4}}$ and $\left|\alpha-\frac{c}{s}\right|\leqq s^{-1}H^{-\frac{9}{4}}Y^{-6}$.

We are given a, q with $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}H^{\frac{3}{4}}Q^{-3}$, $(a,q)=1$, $q\leqq Q^3H^{-\frac{3}{4}}$. Therefore

$$\begin{aligned}
|qc-as|=qs\left|\frac{c}{s}-\frac{a}{q}\right|&\leqq sH^{\frac{3}{4}}Q^{-3}+qH^{-\frac{9}{4}}Y^{-6}\\
&<\frac{1}{2}+Q^3H^{-3}Y^{-6}=\frac{1}{2}+c^{-3}<1.
\end{aligned}$$

Hence $\frac{c}{s}=\frac{a}{q}$ and we finally have the bound

$$P^{\varepsilon}q^{-\frac{1}{3}}\min\left(HY^2, \left|\alpha-\frac{a}{q}\right|^{-\frac{1}{3}}\right)$$

and this is

$$\ll \frac{P^{\varepsilon}HY^2q^{-\frac{1}{3}}}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{3}}}$$

and so is also acceptable. This completes the proof of Lemma 8.

Lemma 9. *Suppose that* $Z<p_1\leqq p_2\leqq 2Z$ *and that* $(a, q)=1$. *Then*

$$g_{p_1p_2}(\alpha)=g^*(\alpha, q, a, p_1, p_2)+O\left(q^{\frac{1}{2}+\varepsilon}\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{2}}\right) \tag{6.7}$$

and

$$h(p_i^3\alpha)=h^*(\alpha, q, a, p_i)+O\left(q^{\frac{1}{2}+\varepsilon}\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{2}}\right) \tag{6.8}$$

where

$$g^*(\alpha, q, a, p_1, p_2)=\zeta(q, a, p_1, p_2)\,J\left(\alpha-\frac{a}{q}, Q\right) \tag{6.9}$$

and

$$h^*(\alpha, q, a, p_i)=\xi(q, a, p_i)\,J\left(p_i^3\left(\alpha-\frac{a}{q}\right), R\right) \tag{6.10}$$

with

$$\xi(q, a, p_1, p_2)=\begin{cases} q^{-1}S(q, a)-(qp_1)^{-1}S(q, ap_1^3) & (p_1=p_2),\\ \dfrac{S(q, a)}{q}-\dfrac{S(q, ap_1^3)}{qp_1}-\dfrac{S(q, ap_2^3)}{qp_2}+\dfrac{S(q, ap_1^3p_2^3)}{qp_1p_2} & (p_1\neq p_2),\end{cases} \tag{6.11}$$

$$\xi(q, a, p_i)=q^{-1}S(q, ap_i^3), \tag{6.12}$$

$$S(q, b)=\sum_{r=1}^{q}e\left(\frac{br^3}{q}\right), \tag{6.13}$$

$$J(\beta, A)=\int_A^{2A}e(\beta x^3)dx. \tag{6.14}$$

We further have

$$g^*(\alpha, q, a, p_1, p_2)\ll \frac{\psi(q)q^{\varepsilon}Q}{1+Q^3\left|\alpha-\frac{a}{q}\right|}, \tag{6.15}$$

$$h^*(\alpha, q, a, p_i) \ll \frac{\psi(q_i) q^{\varepsilon} R}{1 + Q^3 \left|\alpha - \frac{a}{q}\right|} \tag{6.16}$$

where $\psi(q)$ *is the multiplicative function defined by*

$$\psi(p^{3k}) = p^{-k}, \quad \psi(p^{3k+1}) = p^{-k-\frac{1}{2}}, \quad \psi(p^{3k+2}) = p^{-k-1}, \tag{6.17}$$

and

$$q_i = \frac{q}{(q, p_i^3)} \tag{6.18}$$

Proof. The first part of the lemma follows from Theorem 2 of Vaughan [V2]. By Lemmas 4. 3 and 4. 4 of Vaughan [V1], $S(r, b) \ll \psi(r) r^{\varepsilon}$ when $(r, b) = 1$. The second part of the lemma follows from this and by applying partial integration to J.

We now apply the above estimates to T_5.

Let $\mathfrak{n}_1$ denote the set of α in $\mathfrak{n}$ with the property that whenever

$$\left|\alpha - \frac{a}{q}\right| \leqq q^{-1} H^{\frac{7}{4}} Q^{-3}$$

and $(a, q) = 1$ one has $q > H^{\frac{7}{4}}$. Let $\alpha \in \mathfrak{n}_1$. Then, by Lemma 7,

$$\sum_{h \leqq H} |F(\alpha h, \gamma; h)|^2 \ll P^{2+\varepsilon} H^{-\frac{3}{4}}.$$

Choose b, r so that $(r, b) = 1$, $r \leqq Q^3 H^{-\frac{3}{4}}$ and $\left|\alpha - \frac{b}{r}\right| \leqq r^{-1} H^{\frac{3}{4}} Q^{-3}$. Then $\left|\alpha - \frac{b}{r}\right| \leqq r^{-1} H^{\frac{7}{4}} Q^{-3}$, so that $r > H^{\frac{7}{4}} > H^{\frac{3}{4}}$. Therefore, by Lemma 8,

$$\sum_{h \leqq H} |G_h(\alpha h^3, \gamma h)|^2 \ll P^{\varepsilon} H^{\frac{3}{4}} Y^2.$$

Hence, by Cauchy's inequality, (5. 16) and (5. 17),

$$\int_{\mathfrak{n}_1} \sum_{h \leqq H} |F(\alpha h, \gamma; h)\, G_h(\alpha h^3, \theta\gamma h)|\, S(\alpha)\, d\alpha \ll Y Z^4 P^{1+\varepsilon} Q R^2. \tag{6.19}$$

It remains to consider $\mathfrak{n} \setminus \mathfrak{n}_1$. Let

$$\mathfrak{N}_1(q, a) = \left\{\alpha : \left|\alpha - \frac{a}{q}\right| \leqq q^{-1} M Q^{-3}\right\},$$

$$\mathfrak{N}_2(q, a) = \left\{\alpha : \left|\alpha - \frac{a}{q}\right| \leqq q^{-1} H^{\frac{3}{4}} Q^{-3}\right\},$$

$$\mathfrak{N}_3(q, a) = \left\{\alpha : \left|\alpha - \frac{a}{q}\right| \leqq q^{-1} H^{\frac{7}{4}} Q^{-3}\right\},$$

and let $\mathfrak{N}_1$ denote the union of the $\mathfrak{N}_1(q,a)$ with $1 \leqq a \leqq q \leqq M$ and $(a,q)=1$, $\mathfrak{N}_2$ the union of the $\mathfrak{N}_2(q,a)$ with $1 \leqq a \leqq q \leqq H^{\frac{3}{4}}$ and $(a,q)=1$, and $\mathfrak{N}_3$ the union of the $\mathfrak{N}_3(q,a)$ with $1 \leqq a \leqq q \leqq H^{\frac{7}{4}}$ and $(a,q)=1$. Then, modulo 1,

$$\mathfrak{n} = \mathfrak{n}_1 \cup (\mathfrak{N}_3 \setminus \mathfrak{N}_2) \cup (\mathfrak{N}_2 \setminus \mathfrak{N}_1)$$

and, by (6. 19), (5. 27) and (5. 28), it remains to treat $\mathfrak{N}_3 \setminus \mathfrak{N}_2$ and $\mathfrak{N}_2 \setminus \mathfrak{N}_1$.

For brevity write

$$u = u(\alpha, q, a) = \frac{\psi(q) Q}{1 + Q^3 \left|\alpha - \frac{a}{q}\right|},$$

$$v(p) = v(\alpha, q, a, p) = \frac{\psi\left(\frac{q}{(q, p^3)}\right) R}{1 + Q^3 \left|\alpha - \frac{a}{q}\right|},$$

$$v = v(\alpha, q, a) = \frac{\psi(q) R}{1 + Q^3 \left|\alpha - \frac{a}{q}\right|},$$

$$\Delta = \Delta(\alpha, q, a) = q^{\frac{1}{2}} \left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{1}{2}}.$$

By (5. 6) and Lemma 9, when $(a,q)=1$, $q \leqq H^{\frac{7}{4}}$,

$$S(\alpha) \ll P^{\varepsilon}(u^2 + \Delta^2) \Big(\sum_{Z < p \leqq 2Z} (v(p) + \Delta) \Big)^4.$$

Moreover

$$\begin{aligned} \sum_{Z < p \leqq 2Z} \psi\left(\frac{q}{(q, p^3)}\right) &\ll Z\psi(q) + \sum_{\substack{Z < p \leqq 2Z \\ p \| q}} \psi\left(\frac{q}{p}\right) + \sum_{\substack{Z < p \leqq 2Z \\ p^2 \| q}} \psi\left(\frac{q}{p^2}\right) + \sum_{\substack{Z < p \leqq 2Z \\ p^3 \mid q}} \psi\left(\frac{q}{p^3}\right) \\ &\ll Z\psi(q) + \sum_{\substack{Z < p \leqq 2Z \\ p \mid q}} p\psi(q) \end{aligned}$$

so that

$$\sum_{Z < p \leqq 2Z} \psi\left(\frac{q}{(q, p^3)}\right) \ll P^{\varepsilon} Z \psi(q). \tag{6. 20}$$

Therefore

$$S(\alpha) \ll P^{\varepsilon} Z^4 (u^2 + \Delta^2)(v^4 + \Delta^4).$$

Plainly $v < u$. Hence

$$S(\alpha) \ll P^{\varepsilon} Z^4 (u^2 v^4 + u^2 \Delta^4 + \Delta^6). \tag{6.21}$$

Let $\mathfrak{Q}(q, a)$ denote $\mathfrak{N}_3(q, a)$ when $H^{\frac{3}{4}} < q \leqq H^{\frac{7}{4}}$, $1 \leqq a \leqq q$, $(a, q) = 1$ and $\mathfrak{N}_3(q, a) \setminus \mathfrak{N}_2(q, a)$ when $1 \leqq a \leqq q \leqq H^{\frac{3}{4}}$, $(a, q) = 1$. Thus the union of the $\mathfrak{Q}(q, a)$ is $\mathfrak{N}_3 \setminus \mathfrak{N}_2$. Let $\alpha \in \mathfrak{Q}(q, a)$. Then, by Lemma 7,

$$\sum_{h \leqq H} |F(\alpha h, \gamma; h)|^2 \ll \frac{q^{-1} H P^{2+\varepsilon}}{1 + Q^3 \left|\alpha - \frac{a}{q}\right|}. \tag{6.22}$$

Choose b, r so that $\left|\alpha - \frac{b}{r}\right| \leqq r^{-1} H^{\frac{3}{4}} Q^{-3}$, $(r, b) = 1$, $r \leqq Q^3 H^{-\frac{3}{4}}$. If $\frac{b}{r} = \frac{a}{q}$, then by the definition of $\mathfrak{Q}(q, a)$ we have $r > H^{\frac{3}{4}}$. If $\frac{b}{r} \neq \frac{a}{q}$, then

$$\begin{aligned} \frac{1}{rq} \leqq \left|\frac{b}{r} - \frac{a}{q}\right| &\leqq (H^{\frac{3}{4}} r^{-1} + H^{\frac{7}{4}} q^{-1}) Q^{-3} \\ &\leqq \frac{1}{2rq} + H^{\frac{7}{4}} q^{-1} Q^{-3} \end{aligned}$$

and so $r \geqq \frac{1}{2} Q^3 H^{-\frac{7}{4}} > H^{\frac{3}{4}}$ once more. Thus, by Lemma 8,

$$\sum_{h \leqq H} |G_h(\alpha h^3, \gamma h)|^2 \ll Y^2 H^{\frac{3}{4}} P^{\varepsilon},$$

uniformly in γ. Therefore, by (6.22) and Cauchy's inequality

$$\sum_{h \leqq H} |F(\alpha h, \gamma; h)\, G_h(\alpha h^3, \theta\gamma h)| \ll \frac{Y H^{\frac{7}{8}} P^{1+\varepsilon} q^{-\frac{1}{2}}}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{1}{2}}}.$$

We now show that

$$\sum_{q \leqq H^{\frac{7}{4}}} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{\mathfrak{Q}(q,a)} \frac{Y H^{\frac{7}{8}} P q^{-\frac{1}{2}}}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{1}{2}}} S(\alpha)\, d\alpha \ll Y Z^4 P^{1+\varepsilon} Q R^2 \tag{6.23}$$

and this will give a suitable estimate for the contribution from the $\mathfrak{Q}(q, a)$.

By (6. 21) and (6. 22) the integral on the left hand side of (6. 23) is

$$\ll YZ^4H^{\frac{7}{8}}P^{1+\varepsilon}q^{-\frac{1}{2}}\int_{\mathfrak{Q}(q,a)}\left(\frac{Q^2R^4\psi(q)^6}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{13}{2}}}\right.$$
$$\left.+\frac{Q^2\psi(q)^2q^2}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{2}}}+q^3\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{5}{2}}\right)d\alpha$$

and the integral here is

$$\ll Q^{-1}R^4\psi(q)^6\min\left(1,(qH^{-\frac{3}{4}})^{\frac{11}{2}}\right)+Q^2\psi(q)^2q^2Q^{-3}(q^{-1}H^{\frac{7}{4}})^{\frac{1}{2}}+q^3Q^{-3}(q^{-1}H^{\frac{7}{4}})^{\frac{7}{2}}.$$

Therefore, the expression on the left of (6. 23) is

$$\ll YZ^4P^{1+\varepsilon}H^{\frac{7}{8}}\left(Q^{-1}R^4\sum_{q\leqq H^{\frac{7}{4}}}q^{\frac{1}{2}}\psi(q)^6\min\left(1,(qH^{-\frac{3}{4}})^{\frac{11}{2}}\right)\right.$$
$$\left.+H^{\frac{7}{8}}Q^{-1}\sum_{q\leqq H^{\frac{7}{4}}}q^2\psi(q)^2+Q^{-3}(H^{\frac{7}{4}})^{\frac{9}{2}}\right).$$

By writing $q=q_2q_3^3$ where q_2 is cubefree it is easily verified that

$$\sum_{q>\lambda}q^{\frac{1}{2}+\theta}\psi(q)^6\ll\lambda^{\theta-\frac{7}{6}}\quad\text{when}\quad\theta<\frac{7}{6}, \tag{6. 24}$$

$$\sum_{q\geqq\lambda}q^{\frac{1}{2}+\theta}\psi(q)^6\ll\lambda^{\theta-\frac{7}{6}}\quad\text{when}\quad\theta>\frac{7}{6} \tag{6. 25}$$

and

$$\sum_{q\leqq\mu}q^2\psi(q)^2\ll\mu^2. \tag{6. 26}$$

Thus the left hand side of (6. 23) is

$$\ll YZ^4P^{1+\varepsilon}H^{\frac{7}{8}}(Q^{-1}R^4(H^{\frac{3}{4}})^{-\frac{7}{6}}+H^{\frac{35}{8}}Q^{-1}+H^{\frac{63}{8}}Q^{-3})$$
$$\ll YZ^4P^{1+\varepsilon}QR^2$$

as required. The observant reader will notice that it is in the third term here that we are required to have $Y\leqq P^{\frac{17}{115}}$ and it is precisely in this situation that the method of Hooley [H3] is likely to give an improved estimate.

Let $\mathfrak{K}(q,a)$ denote $\mathfrak{N}_2(q,a)$ when $M<q\leqq H^{\frac{3}{4}}$, $1\leqq a\leqq q$, $(a,q)=1$ and $\mathfrak{N}_2(q,a)\setminus\mathfrak{N}_1(q,a)$ when $1\leqq a\leqq q\leqq M$, $(a,q)=1$. Then the union of the $\mathfrak{K}(q,a)$ is $\mathfrak{N}_2\setminus\mathfrak{N}_1$.

Let $\alpha \in \mathfrak{R}(q, a)$. Then, by Lemma 7, (6. 22) holds. Moreover, by Lemma 8,

$$\sum_{h \leqq H} |G_h(\alpha h^3, \gamma h)|^2 \ll \frac{P^\varepsilon Y^2 H q^{-\frac{1}{3}}}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{1}{3}}},$$

uniformly in γ. Therefore, by Cauchy's inequality,

$$\sum_{h \leqq H} |F(\alpha h, \gamma; h)\, G_h(\alpha h^3, \theta\gamma h)| \ll \frac{YHP^{1+\varepsilon} q^{-\frac{2}{3}}}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{2}{3}}}.$$

Hence it suffices to show that

$$\sum_{q \leqq H^{\frac{3}{4}}} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_{\mathfrak{R}(q,a)} \frac{YHPq^{-\frac{2}{3}}}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{2}{3}}} S(\alpha)\, d\alpha \ll YZ^4 P^{1+\varepsilon} QR^2.$$

By (6. 21) the integral on the left is

$$\ll YZ^4 HP^{1+\varepsilon} q^{-\frac{2}{3}} \int_{\mathfrak{R}(q,a)} \left(\frac{Q^2 R^4 \psi(q)^6}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{20}{3}}} + \frac{Q^2 \psi(q)^2 q^2}{\left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{2}{3}}} + q^3 \left(1 + Q^3 \left|\alpha - \frac{a}{q}\right|\right)^{\frac{7}{3}} \right) d\alpha$$

and the integral here is

$$\ll Q^{-1} R^4 \psi(q)^6 \min\left(1, \left(\frac{q}{M}\right)^{\frac{17}{3}}\right) + Q^2 \psi(q)^2\, Q^{-3} (q^{-1} H^{\frac{3}{4}})^{\frac{1}{3}} + q^3 Q^{-3} (q^{-1} H^{\frac{3}{4}})^{\frac{10}{3}}.$$

Therefore the left hand side of (6. 27) is

$$\ll YZ^4 HP^{1+\varepsilon} \left(Q^{-1} R^4 \sum_{q \leqq H^{\frac{3}{4}}} q^{\frac{1}{3}} \psi(q)^6 \min\left(1, \left(\frac{q}{M}\right)^{\frac{17}{3}}\right) + Q^{-1} H^{\frac{1}{4}} \sum_{q \leqq H^{\frac{3}{4}}} q^2 \psi(q)^2 + Q^{-3} H^{\frac{5}{2}} \sum_{q \leqq H^{\frac{3}{4}}} 1 \right).$$

By (6. 24), (6. 25) and (6. 26) this is

$$\begin{aligned} &\ll YZ^4 HP^{1+\varepsilon} (Q^{-1} R^4 M^{-\frac{4}{3}} + Q^{-1} H^{\frac{7}{4}} + Q^{-3} H^{\frac{13}{4}}) \\ &\ll YZ^4 P^{1+\varepsilon} QR^2, \end{aligned}$$

which establishes (6. 27).

This completes the proof of (5. 28).

7. The estimation of T_2

In estimating T_2 it is not so easy to take advantage of the averaging over p. However in compensation the number of terms in our exponential sums is smaller.

By (5. 11) and (5. 12), and Cauchy's inequality,

$$|\Xi_p(\alpha)|^2 \leqq H\Omega_p(\alpha p^2) \tag{7. 1}$$

where

$$\Omega_p(\beta) = \sum_{h \leqq H} \Bigg| \sum_{\substack{\frac{2P}{p} + hp^2 < y \leqq \frac{4P}{p} - hp^2 \\ y \equiv h \pmod 2}} e\left(\frac{3}{4}\beta h y^2\right)\Bigg|^2. \tag{7. 2}$$

Now, in the same manner as Lemma 7, we see that

$$\Omega_p(\beta) \ll \left(\frac{r^{-1}HQ^2}{1 + HQ^2\left|\beta - \frac{b}{r}\right|} + HQ + r\right) P^{\varepsilon} \tag{7. 3}$$

whenever

$$\left|\beta - \frac{b}{r}\right| \leqq r^{-2}, \quad (b, r) = 1. \tag{7. 4}$$

Let $\alpha \in \mathfrak{n}$ and suppose that $Y < p \leqq 2Y$. Choose $b = b(p)$, $r = r(p)$ so that $\left|\alpha p^2 - \frac{b}{r}\right| \leqq r^{-1}HP^{-2}$, $r \leqq P^2H^{-1}$. By (5. 2), $HQ \ll P^2H^{-1}$. Therefore if

$$r\left(1 + HQ^2\left|\alpha p^2 - \frac{b}{r}\right|\right) > H^2Q^2P^{-2},$$

then, by (7. 3) and (7. 1),

$$\Xi_p(\alpha) \ll P^{1+\varepsilon}$$

and so by (5. 15), (5. 16) and (5. 17) the contribution to

$$\sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} T_2(p) \tag{7. 5}$$

is $\ll YZ^4P^{1+\varepsilon}QR^2$. Thus in order to establish (5. 29) it remains to consider the situation when

$$r \leqq H^2Q^2P^{-2} (= H^2Y^{-2}) \quad \text{and} \quad \left|\alpha p^2 - \frac{b}{r}\right| \leqq r^{-1}HP^{-2}.$$

In this case the contribution to (7. 5) from the second and third terms on the right of (7. 3) is also $\ll YZ^4P^{1+\varepsilon}QR^2$. Thus it remains to bound

$$\int_{\mathfrak{n}} \sum_{p\in\mathscr{P}(\alpha)} \frac{P^{\varepsilon}HQ}{\left(r\left(1+HQ^2\left|\alpha p^2-\frac{b}{r}\right|\right)\right)^{\frac{1}{2}}} S(\alpha)\,d\alpha$$

where $\mathscr{P}(\alpha)$ is the set of primes p with $Y<p\leqq 2Y$ for which there are b, r with $(b, r)=1$, $r\leqq H^2Y^{-2}$ and $\left|\alpha p^2-\frac{b}{r}\right|\leqq r^{-1}HP^{-2}$. (Note that if such a fraction $\frac{b}{r}$ exists for a given α and p, then it is unique.)

Given $\alpha\in\mathfrak{n}$, choose a, q so that $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}(8H^2)^{-1}$, $(a, q)=1$, $q\leqq 8H^2$. Then for $p\in\mathscr{P}(\alpha)$ we have

$$\begin{aligned}|bq-arp^2| &= qrp^2\left|\frac{b}{rp^2}-\frac{a}{q}\right| \\ &\leqq HqP^{-2}+rp^2(8H^2)^{-1}<1.\end{aligned}$$

Thus $\frac{b}{rp^2}=\frac{a}{q}$, whence $q=rt$ with $t|p^2$. Now

$$\begin{aligned}\sum_{p\in\mathscr{P}(\alpha)}&\left(r\left(1+HQ^2\left|\alpha p^2-\frac{b}{r}\right|\right)\right)^{-\frac{1}{2}} \\ &\ll\left(q\left(1+HP^2\left|\alpha-\frac{a}{q}\right|\right)\right)^{-\frac{1}{2}}\left(Y+\sum_{\substack{p\\ t(p)=p}} t^{\frac{1}{2}}+\sum_{\substack{p\\ t(p)=p^2}} t^{\frac{1}{2}}\right) \\ &\ll\left(q\left(1+HP^2\left|\alpha-\frac{a}{q}\right|\right)\right)^{-\frac{1}{2}}P^{\varepsilon}Y\end{aligned}$$

since if $t(p)>1$, then $p|q$. If $q\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)>H^2Q^2P^{-2}$, then we have an acceptable bound. Thus we may suppose that $q\leqq H^2Q^2P^{-2}$ and $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}(H^2Q^2P^{-2})Q^{-3}$. By (5. 2),

$$H^2Q^2P^{-2}=H^2Y^2<H^{\frac{7}{4}}.$$

Therefore, in the notation of § 6, $\alpha\in\mathfrak{N}_3$, and since $\alpha\in\mathfrak{n}$, so that $\alpha\notin\mathfrak{N}_1$, we have $\alpha\in\mathfrak{N}_3\setminus\mathfrak{N}_1$. Moreover $HQ<H^{\frac{7}{8}}P$ and, when $q\leqq H^{\frac{3}{4}}$ and $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}H^{\frac{3}{4}}Q^{-3}$ we have

$$\left(q\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)\right)^{\frac{1}{6}}\ll Y.$$

Therefore we are reduced to bounding

$$\sum_{q\leqq H^{\frac{7}{4}}}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\mathfrak{Q}(q,a)}\frac{YH^{\frac{7}{8}}P^{1+\varepsilon}q^{-\frac{1}{2}}}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{1}{2}}}S(\alpha)\,d\alpha$$

and

$$\sum_{q\leqq H^{\frac{3}{4}}}\sum_{\substack{a=1\\(a,q)=1}}^{q}\int_{\mathfrak{R}(q,a)}\frac{YHPq^{-\frac{2}{3}}}{\left(1+Q^3\left|\alpha-\frac{a}{q}\right|\right)^{\frac{2}{3}}}S(\alpha)\,d\alpha$$

and we have obtained suitable estimates for these expressions in (6.23) and (6.27) respectively.

8. A minor arc estimate

Let

$$F(\alpha)=\sum_{\substack{Y<p\leqq 2Y\\p\equiv 2\,(\mathrm{mod}\,3)}}f_p(\alpha)\sum_{\substack{Z<p_1\leqq 2Z\\p_1\equiv 2\,(\mathrm{mod}\,3)}}\sum_{\substack{Z<p_2\leqq 2Z\\p_2\equiv 2\,(\mathrm{mod}\,3)}}g_{p_1p_2}(p^3\alpha)\,h(p^3p_1^3\alpha)\,h(p^3p_2^3\alpha), \tag{8.1}$$

and

$$L=16MY^3=16P^{\frac{78}{115}}. \tag{8.2}$$

We now state the minor arc estimate that is the basis of our proofs of Theorems 4 and 5.

Lemma 10. *Let* $\mathfrak{m}$ *denote the set of* α *in* $(LP^{-3}, 1+LP^{-3}]$ *such that whenever* $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}LP^{-3}$ *and* $(a,q)=1$ *we have* $q>L$. *Then*

$$\int_{\mathfrak{m}}|F(\alpha)|^2\,d\alpha\ll Y^2Z^4P^{1+\varepsilon}QR^2.$$

Proof. Since F is periodic with period 1 we may suppose that $\mathfrak{m}$ is the set of α in $(0,1]$ such that whenever $\left|\alpha-\frac{a}{q}\right|\leqq q^{-1}LP^{-3}$ and $(a,q)=1$ we have $q>L$.

By (8.1), Cauchy's inequality and (5.6) we have

$$\int_{\mathfrak{m}}|F(\alpha)|^2\,d\alpha\ll Y\sum_{\substack{Y<p\leqq 2Y\\p\equiv 2\,(\mathrm{mod}\,3)}}\int_{\mathfrak{m}}|f_p(\alpha)|^2\,S(p^3\alpha)\,d\alpha. \tag{8.3}$$

For a given prime number p with $Y<p\leqq 2Y$ and $p\equiv 2\,(\mathrm{mod}\,3)$ we write

$$d=p^3.$$

Let

$$\mathfrak{n}_k=\{\alpha:\ \alpha-k\in\mathfrak{n}\},$$

$$\mathscr{B}_d=\bigcup_{k=0}^{d-1}\mathfrak{n}_k,$$

$$\mathscr{A}_d=\{\alpha:\ \alpha d\in\mathscr{B}_d\}.$$

We show that $\mathfrak{m} \subset \mathscr{A}_d$. Let $\alpha \in \mathfrak{m}$ and choose k so that $0 \leqq k < \alpha d \leqq k+1 \leqq d$. We show that $\alpha d \in \mathfrak{n}_k$, i.e. that $\alpha d - k \in \mathfrak{n}$. Suppose not. Since $0 < \alpha d - k \leqq 1$ there exist a, q with $(a, q) = 1$, $q \leqq M$, $\left|\alpha d - k - \frac{a}{q}\right| \leqq q^{-1} M Q^{-3}$. Thus $\left|\alpha - \frac{qk+a}{qd}\right| \leqq (qd)^{-1} M Q^{-3}$. Hence there exist b, r with $(b, r) = 1$, $\left|\alpha - \frac{b}{r}\right| \leqq r^{-1} M Q^{-3} \leqq r^{-1} L P^{-3}$ and

$$r \leqq qd \leqq 2M(2Y)^3 = L.$$

This contradicts the definition of $\mathfrak{m}$.

It follows that

$$\int_{\mathfrak{m}} |f_p(\alpha)|^2 \, S(p^3\alpha)\, d\alpha \leqq \int_{\mathscr{A}_d} |f_p(\alpha)|^2 \, S(d\alpha)\, d\alpha. \tag{8.4}$$

By the change of variables $\beta = d\alpha$ we obtain

$$\begin{aligned}\int_{\mathscr{A}_d} |f_p(\alpha)|^2 \, S(d\alpha)\, d\alpha &= \frac{1}{d} \int_{\mathscr{B}_d} \left|f_p\left(\frac{\beta}{d}\right)\right|^2 S(\beta)\, d\beta \\ &= \frac{1}{d} \sum_{k=0}^{d-1} \int_{\mathfrak{n}_k} \left|f_p\left(\frac{\beta}{d}\right)\right|^2 S(\beta)\, d\beta \\ &= \frac{1}{d} \sum_{k=0}^{d-1} \int_{\mathfrak{n}} \left|f_p\left(\frac{\alpha+k}{d}\right)\right|^2 S(\alpha+k)\, d\alpha.\end{aligned}$$

Thus

$$\int_{\mathscr{A}_d} |f_p(\alpha)|^2 \, S(d\alpha)\, d\alpha = \int_{\mathfrak{n}} \frac{1}{d} \sum_{k=0}^{d-1} \left|f_p\left(\frac{\alpha+k}{d}\right)\right|^2 S(\alpha)\, d\alpha. \tag{8.5}$$

By (5.3),

$$\frac{1}{d} \sum_{k=0}^{d-1} \left|f_p\left(\frac{\alpha+k}{d}\right)\right|^2 = \sum_{\substack{P < x_1 \leqq 2P \\ p \nmid x_1 \\ x_1^3 \equiv x_2^3 \,(\mathrm{mod}\, d)}} \sum_{\substack{P < x_2 \leqq 2P \\ p \nmid x_2}} e\left(\frac{\alpha}{d}(x_1^3 - x_2^3)\right).$$

Since $d = p^3$ and $p \equiv 2 \,(\mathrm{mod}\, 3)$, the conditions $p \nmid x_1$, $p \nmid x_2$, $x_1^3 \equiv x_2^3 \,(\mathrm{mod}\, d)$ are equivalent to $p \nmid x_1$, $p \nmid x_2$, $x_1 \equiv x_2 \,(\mathrm{mod}\, d)$. Let $h = \frac{x_2 - x_1}{d}$, $y = x_2 + x_1$. Then the summation conditions are equivalent to $2P < y + hp^3 \leqq 4P$, $2P < y - hp^3 \leqq 4P$, $p \nmid y$, $y \equiv h \,(\mathrm{mod}\, 2)$. Thus the double sum becomes

$$\sum_h \sum_{\substack{P + |h|p^3 < y \leqq 2P - |h|p^3 \\ p \nmid y,\, y \equiv h \,(\mathrm{mod}\, 2)}} e\left(\frac{3\alpha}{4d} hp^3 y^2 + \frac{\alpha}{4d} h^3 p^9\right)$$

$$= \sum_{\substack{P < y \leqq 2P \\ p \nmid y}} 1 + 2\,\mathrm{Re} \sum_{h > 0} \sum_{\substack{P + hp^3 < y \leqq 2P - hp^3 \\ p \nmid y,\, y \equiv h \,(\mathrm{mod}\, 2)}} e\left(\frac{3}{4}\alpha h y^2 + \frac{1}{4}\alpha h^3 p^6\right).$$

Clearly the innermost summation conditions imply that $h \leqq H$. Therefore, by (5. 7),

$$\frac{1}{d} \sum_{k=0}^{d-1} \left| f_p\left(\frac{\alpha+k}{d}\right)\right|^2 = \Phi_p(\alpha).$$

Thus, by (8. 3), (8. 4) and (8. 5),

$$\int_{\mathfrak{m}} |F(\alpha)|^2 d\alpha \ll Y \int_{\mathfrak{n}} \sum_{\substack{Y<p\leqq 2Y \\ p\equiv 2 \,(\mathrm{mod}\, 3)}} \Phi_p(\alpha)\, S(\alpha) d\alpha.$$

Lemma 10 now follows from Lemma 6.

9. The proof of Theorem 4

Let

$$P = \left[\frac{1}{3} N^{\frac{1}{3}}\right], \tag{9. 1}$$

let $R(n)$ denote the number of choices of p, p_1, p_2, x, y, z_1, z_2 such that

$$x^3 + p^3 y^3 + p^3 p_1^3 z_1^3 + p^3 p_2^3 z_2^3 = n, \tag{9. 2}$$

$$P < x \leqq 2P, \quad Q < y \leqq 2Q, \quad R < z_i \leqq 2R, \tag{9. 3}$$

$$Y < p \leqq 2Y, \quad Z < p_i \leqq 2Z, \quad p \equiv 2 \,(\mathrm{mod}\, 3), \quad p_i \equiv 2 \,(\mathrm{mod}\, 3) \tag{9. 4}$$

and

$$p \nmid x, \quad p_i \nmid y, \tag{9. 5}$$

and let

$$R_{\mathscr{A}}(n) = \int_{\mathscr{A}} F(\alpha)\, e(-\alpha n) d\alpha \tag{9. 6}$$

where F is given by (8. 1). Then $R(n) = R_{(0,1]}(n)$.

When $1 \leqq a \leqq q \leqq L$ and $(a, q) = 1$ we take $\mathfrak{M}(q, a)$ to be the interval $\left\{\alpha: \left|\alpha - \frac{a}{q}\right| \leqq q^{-1} L P^{-3}\right\}$ and let $\mathfrak{M}$ denote the union of all such $\mathfrak{M}(q, a)$. Clearly $\mathfrak{M}$ and $\mathfrak{m}$ are disjoint and their union is $(LP^{-3}, 1 + LP^{-3}]$. Moreover

$$R(n) = R_{\mathfrak{M}}(n) + R_{\mathfrak{m}}(n). \tag{9. 7}$$

Now $R_{\mathfrak{m}}(n)$ is the Fourier coefficient of the function which is $F(\alpha)$ on $\mathfrak{m}$ and 0 on $\mathfrak{M}$. Thus, by Bessel's inequality,

$$\sum_{n \leqq N} |R_{\mathfrak{m}}(n)|^2 \leqq \int_{\mathfrak{m}} |F(\alpha)|^2 d\alpha.$$

Therefore, by Lemma 10,

$$\sum_{n \leqq N} |R_{\mathfrak{m}}(n)|^2 \ll (PY^{-2})^2 N^{\frac{103}{115}+\varepsilon}, \tag{9.8}$$

since $P^5 Q^{-1} R^2 \ll N^{\frac{103}{115}}$.

We now proceed to estimate $R_{\mathfrak{M}}(n)$. The argument is similar, to start with, to the estimation of the contribution from the major arcs in Lemma 6. Let $\alpha \in \mathfrak{M}(q, a)$. Then, by Theorem 4.1 of Vaughan [V1], and in the notation of Lemma 9,

$$f_p(\alpha) = f^*(\alpha, q, a, p) + \Delta_f$$

where

$$f^*(\alpha, q, a, p) = \eta(q, a, p)\, J\left(\alpha - \frac{a}{q}, P\right) \tag{9.9}$$

with

$$\eta(q, a, p) = q^{-1} S(q, a) - q^{-1} p^{-1} S(q, ap^3), \tag{9.10}$$

and

$$\Delta_f \ll \Delta$$

where

$$\Delta = q^{\frac{1}{2}+\varepsilon}. \tag{9.11}$$

We further have

$$g_{p_1 p_2}(p^3\alpha) = g^*(p^3\alpha, q, ap^3, p_1, p_2) + \Delta_g,$$
$$h(p^3 p_i^3 \alpha) = h^*(p^3\alpha, q, ap^3, p_i) + \Delta_i$$

with

$$\Delta_g \ll \Delta, \quad \Delta_i \ll \Delta.$$

For brevity write

$$f^* = f^*(\alpha, q, a, p),$$
$$g^* = g^*(p^3\alpha, q, ap^3, p_1, p_2),$$
$$h_i^* = h^*(p^3\alpha, q, ap^3, p_i).$$

It follows that for $\alpha \in \mathfrak{M}(q, a)$,

$$f^* \ll tP^{\varepsilon}, \quad g^* \ll uP^{\varepsilon}, \quad h_i^* \ll v_i P^{\varepsilon}$$

where

$$t = \frac{\psi(q) P}{1 + P^3 \left|\alpha - \frac{a}{q}\right|}, \tag{9.12}$$

$$u=\frac{\psi(q_0)Q}{1+P^3\left|\alpha-\frac{a}{q}\right|},\quad v_i=\frac{\psi(q_i)R}{1+P^3\left|\alpha-\frac{a}{q}\right|} \tag{9.13}$$

with

$$q_0=\left(\frac{q}{(q,p^3)}\right),\quad q_i=\left(\frac{q_0}{(q_0,p_i^3)}\right).$$

We also have

$$\sum_{Z<p_i\leqq 2Z}\psi(q_i)\ll ZP^{\varepsilon}\psi(q_0)$$

so that

$$\sum_{Z<p_i\leqq 2Z}v_i\ll ZP^{\varepsilon}v$$

with

$$v=\frac{\psi(q_0)R}{1+P^3\left|\alpha-\frac{a}{q}\right|}. \tag{9.14}$$

Since $\alpha\in\mathfrak{M}(q,a)$, we have

$$\Delta<v<u<t.$$

Therefore, on writing

$$F^*(\alpha,q,a)=\sum_{\substack{Y<p\leqq 2Y\\ p\equiv 2\,(\mathrm{mod}\,3)}}\ \sum_{\substack{Z<p_1\leqq 2Z\\ p_1\equiv 2\,(\mathrm{mod}\,3)}}\ \sum_{\substack{Z<p_2\leqq 2Z\\ p_2\equiv 2\,(\mathrm{mod}\,3)}} f^*g^*h_1^*h_2^*, \tag{9.15}$$

$$D(\alpha,q,a)=F(\alpha)-F^*(\alpha,q,a)$$

we see that

$$D(\alpha,q,a)\ll\sum_{\substack{Y<p\leqq 2Y\\ p\equiv 2\,(\mathrm{mod}\,3)}}Z^2P^{\varepsilon}tuv\Delta. \tag{9.16}$$

We define $F^*(\alpha)$, $D(\alpha)$ on $(LP^{-3},1+LP^{-3}]$ by

$$F^*(\alpha)=\begin{cases}F^*(\alpha,q,a) & \alpha\in\mathfrak{M}(q,a),\\ 0 & \alpha\notin\mathfrak{M},\end{cases} \tag{9.17}$$

$$D(\alpha)=\begin{cases}D^*(\alpha,q,a) & \alpha\in\mathfrak{M}(q,a),\\ 0 & \alpha\notin\mathfrak{M}\end{cases} \tag{9.18}$$

and let

$$R_0(n)=\int_{\mathfrak{M}}F^*(\alpha)\,e(-\alpha n)\,d\alpha. \tag{9.19}$$

Then

$$R_{\mathfrak{M}}(n) - R_0(n) = \int_{\mathfrak{M}} D(\alpha)\, e(-\alpha n)\, d\alpha$$

and, by Bessel's inequality,

$$\sum_{n \leqq N} |R_{\mathfrak{M}}(n) - R_0(n)|^2 \leqq \int_{\mathfrak{M}} |D(\alpha)|^2\, d\alpha. \tag{9.20}$$

By (9.16), (9.11), (9.12), (9.13) and (9.14),

$$\int_{\mathfrak{M}} |D(\alpha)|^2\, d\alpha \ll Z^4 P^{\varepsilon} \sum_{Y < p \leqq 2Y} \sum_{Y < p' \leqq 2Y} \sum_{q \leqq L} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} \int_0^{q^{-1}LP^{-3}} \frac{P^2 Q^2 R^2 \psi(q)^2\, \psi(q_0)^2\, \psi(q_0')^2\, q}{(1+P^3\beta)^6}\, d\beta$$

where

$$q_0' = \frac{q}{(q, p'^3)}.$$

Hence

$$\int_{\mathfrak{M}} |D(\alpha)|^2\, d\alpha \ll Z^4 P^{\varepsilon - 1} Q^2 R^2 \sum_{Y < p \leqq 2Y} \sum_{Y < p' \leqq 2Y} \sum_{q \leqq L} q^2 \psi(q)^2\, \psi(q_0)^2\, \psi(q_0')^2.$$

When $p \neq p'$, the general term in the innermost sum is of the form

$$p^{2k} \psi(p^k)^4\, \psi\left(\frac{p^k}{(p^k, p^3)}\right)^2 (p')^{2l} \psi(p'^l)^4\, \psi\left(\frac{p'^l}{(p'^l, p'^3)}\right)^2 r^2 \psi(r)^6$$

with $(pp', r) = 1$, $p^k p'^l r \leqq L$. Each r can be written uniquely in the form $r_2 r_3^3$ where r_2 is cubefree. Thus summing over r gives a contribution

$$\leqq \sum_{r_2 r_3^3 \leqq L p^{-k} (p')^{-l}} r_2^{-1} \ll P^{\varepsilon} \left(L p^{-k} (p')^{-l}\right)^{\frac{1}{3}}.$$

Moreover $0 \leqq k$, $0 \leqq l$, and $k + l \leqq 4$. It follows that the sum over k, l, r contributes

$$\ll P^{\varepsilon} L^{\frac{1}{3}} p \ll P^{\varepsilon} L^{\frac{1}{3}} Y.$$

When $p = p'$ the general term in the innermost sum is of the form

$$p^{2k} \psi(p^k)^2\, \psi\left(\frac{p^k}{(p^k, p^3)}\right)^4 r^2 \psi(r)^6$$

with $(p, r) = 1$, $p^k r \leqq L$. Summing over r gives a contribution

$$\ll P^{\varepsilon} (L p^{-k})^{\frac{1}{3}}.$$

Moreover $0 \leqq k \leqq 4$ and the largest term occurs when $k=3$. Thus the sum over k, r contributes

$$\ll P^{\varepsilon} L^{\frac{1}{3}} Y^3.$$

Therefore

$$\int_{\mathfrak{M}} |D(\alpha)|^2 d\alpha \ll Z^4 P^{\varepsilon-1} Q^2 R^2 L^{\frac{1}{3}} Y^4 = (YZ^2 P^{-2} Q R^2)^2 P^{3+\varepsilon} R^{-2} L^{\frac{1}{3}} Y^2.$$

Hence, by (9. 19),

$$\sum_{n \leqq N} |R_{\mathfrak{M}}(n) - R_0(n)|^2 \ll (PY^{-2})^2 N^{\frac{103}{115}}. \tag{9. 21}$$

Let

$$\mathfrak{M}_0(q, a) = \left\{\alpha: q^{-1} LP^{-3} < \left|\alpha - \frac{a}{q}\right| \leqq \frac{1}{3} L^{-2}\right\}. \tag{9. 22}$$

Then the $\mathfrak{M}_0(q, a)$ with $1 \leqq a \leqq q \leqq L$ and $(a, q)=1$ are disjoint and contained in $\left(\frac{1}{3} L^{-2}, 1+\frac{1}{3} L^{-2}\right]$. Let $\mathfrak{M}_0$ denote the union of these $\mathfrak{M}_0(q, a)$, let

$$\mathfrak{M}_1(q, a) = \mathfrak{M}(q, a) \cup \mathfrak{M}_0(q, a), \quad \mathfrak{M}_1 = \mathfrak{M} \cup \mathfrak{M}_0. \tag{9. 23}$$

For α in $\left(\frac{1}{3} L^{-2}, 1+\frac{1}{3} L^{-2}\right]$ we define $F_0^*(\alpha)$, $F_1^*(\alpha)$ by

$$F_i^*(\alpha) = \begin{cases} F^*(\alpha, q, a) & \alpha \in \mathfrak{M}_i(q, a) \\ 0 & \alpha \notin \mathfrak{M}_i \end{cases}. \tag{9. 24}$$

Let

$$R_1(n) = \int_{\mathfrak{M}_1} F_1^*(\alpha)\, e(-\alpha n) d\alpha. \tag{9. 25}$$

Then, by (9. 17) and (9. 19),

$$R_1(n) - R_0(n) = \int_{\frac{1}{3} L^{-2}}^{1+\frac{1}{3} L^{-2}} F_0^*(\alpha)\, e(-\alpha n) d\alpha.$$

Therefore, by Bessel's inequality,

$$\sum_{n \leqq N} |R_1(n) - R_0(n)|^2 \leqq \int_{\mathfrak{M}_0} |F_0^*(\alpha)|^2 d\alpha. \tag{9. 26}$$

Proceeding as above we find that

$$\int_{\mathfrak{M}_0} |F_0^*(\alpha)|^2 d\alpha \ll Z^4 P^{\varepsilon-1} Q^2 R^4 \sum_{Y < p \leqq 2Y} \sum_{Y < p' \leqq 2Y} \sum_{q \leqq L} \left(\frac{q}{L}\right)^7 q \psi(q)^2\, \psi(q_0)^3\, \psi(q_0')^3.$$

Now

$$\sum_{r \leqq \lambda} r^8 \psi(r)^8 = \sum_{r_2 r_3^3 \leqq \lambda} r_2^4 r_3^{16} \ll \lambda^{\frac{17}{3}}.$$

Thus when $p \neq p'$ the contribution from the innermost sum is

$$\ll L^{-\frac{4}{3}} \sum_{\substack{k \geqq 0 \\ k+l \leqq 4}} \sum_{l \geqq 0} p^{\frac{7k}{3}} \psi(p^k)^5 \psi\left(\frac{p^k}{(p^k, p^3)}\right)^3 (p')^{\frac{7l}{3}} \psi(p'^l)^5 \psi\left(\frac{p'^l}{(p'^l, p'^3)}\right)^3$$

$$\ll L^{-\frac{4}{3}}(p^2 + p'^2) \ll L^{-\frac{4}{3}} Y^2,$$

and when $p = p'$ it is

$$\ll L^{-\frac{4}{3}} \sum_{k=0}^{4} p^{\frac{7k}{3}} \psi(p^k)^2 \psi\left(\frac{p^k}{(p^k, p^3)}\right)^6$$

$$\ll L^{-\frac{4}{3}} p^5 \ll L^{-\frac{4}{3}} Y^5.$$

Therefore

$$\int_{\mathfrak{M}_0} |F^*(\alpha)|^2 \, d\alpha \ll Y^2 Z^4 P^{\varepsilon - 1} Q^2 R^4 L^{-\frac{4}{3}} Y^4.$$

Hence, by (9. 26),

$$\sum_{n \leqq N} |R_1(n) - R_0(n)|^2 \ll (PY^{-2})^2 N^{\frac{103}{115} + \varepsilon}. \tag{9. 27}$$

By (9. 25), (9. 22), (9. 23), (9. 15), (9. 9), (6. 9) and (6. 10),

$$R_1(n) = \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \pmod 3}} \sum_{\substack{Z < p_1 \leqq 2Z \\ p_1 \equiv 2 \pmod 3}} \sum_{\substack{Z < p_2 \leqq 2Z \\ p_2 \equiv 2 \pmod 3}} \mathfrak{S}(n, p, p_1, p_2, L) \, I(n, p, p_1, p_2) \tag{9. 28}$$

where

$$\mathfrak{S}(n, p, p_1, p_2, L) = \sum_{q \leqq L} T(q), \tag{9. 29}$$

$$T(q) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} T(q, a) \, e\left(-\frac{an}{q}\right), \tag{9. 30}$$

$$T(q, a) = \eta(q, a, p) \, \zeta(q, ap^3, p_1, p_2) \, \xi(q, ap^3, p_1) \, \xi(q, ap^3, p_2), \tag{9. 31}$$

$$I(n, p, p_1, p_2) = \int_{-\frac{1}{3L^2}}^{\frac{1}{3L^2}} J(\beta, P) \, J(\beta p^3, Q) \, J(\beta p^3 p_1^3, R) \, J(\beta p^3 p_2^3, R) \, e(-\beta n) \, d\beta. \tag{9. 32}$$

It is easily verified that

$$I(n, p, p_1, p_2) = I_1(n, p, p_1, p_2) + O(PQR^2 P^{-12} L^6) \tag{9. 33}$$

where

$$I_1(n, p, p_1, p_2) = \int_{-\infty}^{\infty} J(\beta, P) \, J(\beta p^3, Q) \, J(\beta p^3 p_1^3, R) \, J(\beta p^3 p_2^3, R) \, e(-\beta n) \, d\beta.$$

By a change of variables in each J we obtain

$$I_1(n, p, p_1, p_2) = \frac{1}{p^3 p_1 p_2} \int_{-\infty}^{\infty} K(\beta, P)\, K(\beta, pQ)\, K(\beta, pp_1 R)\, K(\beta, pp_2 R)\, e(-\beta n) d\beta$$

where

$$K(\beta, A) = \int_{A^3}^{8A^3} \frac{1}{3} y^{-\frac{2}{3}} e(\beta y) dy.$$

Let $\mathscr{R}(v)$ denote the set of points $(\alpha_1, \alpha_2, \alpha_3)$ in $\mathbb{R}^3$ with $P^3 \leqq \alpha_1 \leqq 8P$, $p^3 Q^3 \leqq \alpha_2 \leqq 8p^3 Q^3$, $p^3 p_1^3 R^3 \leqq \alpha_3 \leqq 8p^3 p_1^3 R^3$, $p^3 p_2^3 R^3 \leqq v - \alpha_1 - \alpha_2 - \alpha_3 \leqq 8p^3 p_2^3 R^3$. Then, by Fubini's theorem,

$$K(\beta, P)\, K(\beta, pQ)\, K(\beta, pp_1 R)\, K(\beta, pp_2 R) = \hat{V}(\beta)$$

where

$$\hat{V}(\beta) = \int_{-\infty}^{\infty} e(\beta v)\, V(v) dv$$

and

$$V(v) = \int_{\mathscr{R}(v)} \frac{1}{81} \left(\alpha_1 \alpha_2 \alpha_3 (v - \alpha_1 - \alpha_2 - \alpha_3)\right)^{-\frac{2}{3}} d\alpha_1\, d\alpha_2\, d\alpha_3.$$

The function $V(v)$ is continuous and has compact support. Hence, by the Fourier inversion theorem,

$$I_1(n, p, p_1, p_2) = \frac{1}{p^3 p_1 p_2} \int_{-\infty}^{\infty} \hat{V}(\beta)\, e(-\beta n) d\beta$$
$$= \frac{V(n)}{p^3 p_1 p_2}.$$

Thus

$$0 \leqq I_1(n, p, p_1, p_2) \ll P Y^{-3} Z^{-2}. \tag{9.34}$$

Moreover, it follows from our choice of P, (9.1), that if δ is a sufficiently small positive constant and if

$$Y < p \leqq Y(1+\delta), \tag{9.35}$$

$$Z < p_i \leqq Z(1+\delta) \tag{9.36}$$

and

$$\frac{1}{2} N < n \leqq N, \tag{9.37}$$

then

$$I_1(n, p, p_1, p_2) \gg P Y^{-3} Z^{-2}. \tag{9.38}$$

We now turn our attention to $\mathfrak{S}(n, p, p_1, p_2, L)$. By (9. 31), (9. 10), (6. 11) and (6. 12),

$$T(q, a) = q^{-4}\theta(q, p, p_1, p_2)\, U(q, a) \tag{9. 39}$$

with

$$U(q, a) = \sum_{x, y, z_1, z_2} e\left(\frac{a}{q}(x^3 + p^3y^3 + p^3p_1^3z_1^3 + p^3p_2^3z_2^3)\right)$$

and the summation variables subject to

$$1 \leqq x \leqq q, \quad 1 \leqq y \leqq q, \quad 1 \leqq z_i \leqq q, \quad (x, q, p) = 1, \quad (y, q, p_1p_2) = 1,$$

and with

$$\theta(q, p, p_1, p_2) = \prod_{\substack{\pi \in \{p, p_1, p_2\} \\ \pi \nmid q}} \left(1 - \frac{1}{\pi}\right). \tag{9. 40}$$

Suppose for the time being that $p_1 \neq p_2$. Then by a standard argument (see, for example, § 2. 6 of Vaughan [V1]),

$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} U(q, a)\, e\left(-\frac{an}{q}\right) = S(q_0)\, S^*(p^t)\, S_1(p_1^{t_1})\, S_2(p_2^{t_2})$$

where

$$p^t \| q, \quad p_i^{t_i} \| q, \quad q = p^t p_1^{t_1} p_2^{t_2} q_0,$$

$$S(r) = \sum_{\substack{a=1 \\ (a,r)=1}}^{r} S(r, a)^4\, e\left(-\frac{an}{r}\right),$$

$$S^*(r) = \sum_{\substack{a=1 \\ (a,r)=1}}^{r} W(r, a)\, S(r, p^3a)^3\, e\left(-\frac{an}{r}\right), \tag{9. 42}$$

$$S_i(r) = \sum_{\substack{a=1 \\ (a,r)=1}}^{r} S(r, a)^2\, W(r, a)\, S(r, p_i^3a)\, e\left(-\frac{an}{r}\right)$$

and

$$W(r, a) = \sum_{\substack{x=1 \\ (x,r)=1}}^{r} e\left(\frac{ax^3}{r}\right).$$

For a prime number $\pi \equiv 2 \pmod 3$ we have

$$W(\pi^\tau, a) = 0$$

when $\pi \nmid a$ and $\tau \geqq 2$, and

$$S(\pi, a) = 0$$

when $\pi \nmid a$. Thus

$$S^*(p^t) = \begin{cases} 0 & \text{when} \quad t \geqq 2, \\ p^3 & \text{when} \quad t = 1 \quad \text{and} \quad p \nmid n, \\ p^3 - p^4 & \text{when} \quad t = 1 \quad \text{and} \quad p \mid n \end{cases} \tag{9.43}$$

and

$$S_i(p_i^{t_i}) = 0 \quad \text{when} \quad t_i \geqq 1. \tag{9.44}$$

Therefore, by (9.39), (9.40) and Lemma 4.7 of Vaughan [V1] we have

$$T(q) \ll (q, n)\,(q, p)^{\frac{1}{3}} q^{\varepsilon - \frac{4}{3}}. \tag{9.45}$$

When $p_1 = p_2$ the above has to be modified as follows. The equation (9.41) becomes, with the above notation,

$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} U(q, a)\, e\left(-\frac{an}{q}\right) = S(q_0)\, S^*(p^t)\, S_3(p_1^{t_1}) \tag{9.46}$$

where

$$S_3(r) = \sum_{\substack{a=1 \\ (a,r)=1}}^{r} S(r, a)\, W(r, a)\, S(r, p_1^3 a)^2\, e\left(-\frac{an}{r}\right).$$

In the same manner as for S_i we have

$$S_3(p_1^{t_1}) = 0 \quad \text{when} \quad t_1 \geqq 1. \tag{9.47}$$

Thus (9.45) holds once more.

By (9.45),

$$\sum_{q > L} T(q) \ll P^{\varepsilon} \left(\frac{p}{L}\right)^{\frac{1}{3}} \ll P^{-\frac{1}{6}} \tag{9.48}$$

and

$$\mathfrak{S}(n, p, p_1, p_2) \ll P^{\varepsilon} \tag{9.49}$$

where

$$\mathfrak{S}(n, p, p_1, p_2) = \sum_{q=1}^{\infty} T(q), \tag{9.50}$$

and the infinite series in (9.50) converges absolutely. Hence, by (9.28) and (9.34),

$$R_1(n) = R_2(n) + O(P^{\frac{5}{6}} Y^{-2}) \tag{9.51}$$

where

(9.52)

$$R_2(n) = \sum_{\substack{Y < p \leqq 2Y \\ p \equiv 2 \,(\mathrm{mod}\, 3)}} \sum_{\substack{Z < p_1 \leqq 2Z \\ p_1 \equiv 2 \,(\mathrm{mod}\, 3)}} \sum_{\substack{Z < p_2 \leqq 2Z \\ p_2 \equiv 2 \,(\mathrm{mod}\, 3)}} \mathfrak{S}(n, p, p_1, p_2)\, I(n, p, p_1, p_2).$$

By (9. 42) and Lemma 2. 11 of Vaughan [V1], $S(q)$ is a multiplicative function. Suppose that $p_1 \neq p_2$. Then, by (9. 39), (9. 41) and (9. 50)

$$\mathfrak{S}(n, p, p_1, p_2) = \mathfrak{S}_0 \mathfrak{S}_1 \mathfrak{S}_2 \prod_{\pi \notin \{p, p_1, p_2\}} \left(\sum_{k=0}^{\infty} \pi^{-4k} S(\pi^k) \right) \tag{9.53}$$

where

$$\mathfrak{S}_0 = 1 - \frac{1}{p} + \sum_{k=1}^{\infty} p^{-4k} S^*(p^k)$$

and

$$\mathfrak{S}_i = 1 - \frac{1}{p_i} + \sum_{k=1}^{\infty} p_i^{-4k} S_i(p_i^k).$$

By the theory of the singular series in Waring's problem (see §§ 4. 3 and 4. 5 of Vaughan [V1]) the infinite product in (9. 53) is $\gg 1$. Moreover, by (9. 43) and (9. 44),

$$\mathfrak{S}_0 = \begin{cases} 1 & \text{when } p \nmid n, \\ 0 & \text{when } p \mid n, \end{cases}$$

$$\mathfrak{S}_i = 1 - \frac{1}{p_i}.$$

When $p_1 = p_2$, we have to replace $\mathfrak{S}_1 \mathfrak{S}_2$ by

$$\mathfrak{S}_3 = 1 - \frac{1}{p_1} + \sum_{k=1}^{\infty} p_1^{-4k} S_3(p_1^k).$$

By (9. 47), $\mathfrak{S}_3 = 1 - \frac{1}{p_1}$. Hence we see that

$$\mathfrak{S}(n, p, p_1, p_2) \gg 1 \quad \text{when} \quad p \nmid n,$$
$$\mathfrak{S}(n, p, p_1, p_2) = 0 \quad \text{when} \quad p \mid n.$$

Therefore, by (9. 52) and (9. 38),

$$R_2(n) \gg PY^{-2}(\log P)^{-3} \quad \left(\frac{1}{2} N < n \leqq N\right). \tag{9.54}$$

To summarize, by (9. 7), (9. 8), (9. 21) and (9. 27) we have

$$\sum_{n \leqq N} |R(n) - R_1(n)|^2 \ll (PY^{-2}(\log P)^{-3})^2 N^{\frac{103}{115} + \varepsilon} \tag{9.55}$$

and, by (9. 51) and (9. 54),

$$R_1(n) \gg PY^{-2}(\log P)^{-3} \quad \left(\frac{1}{2} N < n \leqq N\right). \tag{9.56}$$

Theorem 4 follows easily from this.

10. The proof of Theorem 5

Let $R(N)$ and $R_1(N)$ be as in § 9, let

$$X=\frac{1}{10}N^{\frac{1}{3}}, \tag{10.1}$$

and let $r(n)$ denote the number of choices of x, p_1, p_2, y_1, y_2 with

$$x\leqq X,\quad X^{\frac{1}{7}}<p_i\leqq 2X^{\frac{1}{7}},\quad y_i\leqq X^{\frac{6}{7}},\quad p_i\nmid x$$

and

$$x^3+p_1^3y_1^3+p_2^3y_2^3=n.$$

Let

$$f_d(\alpha)=\sum_{\substack{x\leqq X\\(x,d)=1}}e(\alpha x^3),\quad g(\alpha)=\sum_{y\leqq X^{\frac{6}{7}}}e(\alpha y^3).$$

Then

$$\sum_n r(n)^2=\int_0^1\Bigg|\sum_{X^{\frac{1}{7}}<p_1\leqq 2X^{\frac{1}{7}}}\ \sum_{X^{\frac{1}{7}}<p_2\leqq 2X^{\frac{1}{7}}}f_{p_1p_2}(\alpha)\,g(p_1^3\alpha)\,g(p_2^3\alpha)\Bigg|^2d\alpha.$$

From the inequality $|zw|\leqq\frac{1}{2}|z|^2+\frac{1}{2}|w|^2$ and on interchanging p_1 and p_2 we obtain

$$\sum_n r(n)^2\leqq\int_0^1\Bigg(\sum_{X^{\frac{1}{7}}<p_1\leqq 2X^{\frac{1}{7}}}\ \sum_{X^{\frac{1}{7}}<p_2\leqq 2X^{\frac{1}{7}}}|f_{p_1p_2}(\alpha)\,g(p_1^3\alpha)^2|\Bigg)^2d\alpha.$$

Hence, by Cauchy's inequality,

$$\sum_n r(n)^2\ll X^{\frac{2}{7}}\int_0^1\sum_{X^{\frac{1}{7}}<p\leqq 2X^{\frac{1}{7}}}\ \sum_{X^{\frac{1}{7}}<p\leqq 2X^{\frac{1}{7}}}|f_{p_1p_2}(\alpha)^2\,g(p_1^3\alpha)^4|d\alpha.$$

Thus, by an appeal to Theorem A we obtain

$$\sum_n r(n)^2\ll X^{\frac{23}{7}+\varepsilon}\ll N^{\frac{23}{21}+\varepsilon}.$$

We also have

$$\sum_{n\leqq\frac{1}{3}N}r(n)\gg N(\log N)^{-2}.$$

Consider

$$\rho(N)=\sum_{\frac{1}{2}N<n<N}R(n)\,r(N-n).$$

Then

$$\rho(N)=\sum_{\frac{1}{2}N<n<N}R_1(n)\,r(N-n)+\sum_{\frac{1}{2}N<n<N}\big(R(n)-R_1(n)\big)\,r(N-n).$$

By (9. 56) the first sum is

$$\gg PY^{-2}(\log P)^{-3} \sum_{m \leqq \frac{1}{3}N} r(m) \gg PY^{-2}N(\log N)^{-5}$$

and by (9. 55) and Cauchy's inequality the second sum is

$$\ll PY^{-2}N^{\frac{103}{230}+\frac{23}{42}+\varepsilon} = PY^{-2}N^{\frac{2404}{2415}+\varepsilon}.$$

Therefore, by (5. 2) and (9. 1),

$$\rho(N) \gg N^{\frac{4}{3}-\frac{34}{345}}(\log N)^{-5}.$$

Moreover, since for any fixed $\delta > 0$, a number not exceeding N can have at most $O(1)$ prime divisors greater than N^{δ}, it follows that

$$\mathscr{R}_7(N) \gg \rho(N) \gg N^{\frac{4}{3}-\frac{34}{345}}(\log N)^{-5}$$

which completes the proof of Theorem 5.

References

[D1] *H. Davenport*, On Waring's problem for cubes, Acta Math. **71** (1939), 123—143.

[D2] *H. Davenport*, Sums of three positive cubes, J. London Math. Soc. **25** (1950), 339—343.

[Ha] *H. Halberstam* and *H.-E. Richert*, Sieve Methods, London 1974.

[H1] *C. Hooley*, On a new technique and its applications to the theory of numbers, Proc. London Math. Soc. (3) **38** (1979), 115—151.

[H2] *C. Hooley*, On the numbers that are representable as the sum of two cubes, J. reine angew. Math. **314** (1980), 146—173.

[H3] *C. Hooley*, On Waring's problem for cubes, Acta Math. to appear.

[L] *Yu. V. Linnik*, On the representation of large numbers as sums of seven cubes, Mat. Sbornik **12** (1943), 218—224.

[T] *E. C. Titchmarsh*, The Theory of the Riemann zeta-function, Oxford 1951.

[V1] *R. C. Vaughan*, The Hardy-Littlewood Method, London 1981.

[V2] *R. C. Vaughan*, Some remarks on Weyl sums, Coll. Math. Soc. János Bolyai, Budapest 1981.

[V3] *R. C. Vaughan*, Sums of three cubes, Bull. London Math. Soc. **17** (1985), 17—20.

[W] *G. L. Watson*, A proof of the seven cube theorem, J. London Math. Soc. **26** (1951), 153—156.

Imperial College, London, SW7 2BZ, England

Eingegangen 24. Mai 1985

On divisors of sums of integers · II

By *A. Sárközy* at Budapest and *C. L. Stewart**) at Waterloo

1. Introduction

Throughout this article, $c_0, c_1, c_2, \ldots$ will denote effectively computable positive absolute constants. Denote the cardinality of a set X by $|X|$ and for any integer n let $P(n)$ denote the greatest prime factor of n with the convention that $P(0)=P(\pm 1)=1$. Let N be a positive integer and let A and B be non-empty subsets of $\{1,\ldots,N\}$. In [2] Balog and Sárközy proved, by means of the large sieve, that if $N>N_0$, and $|A|\,|B|>100N(\log N)^2$ then there exist $a\in A$ and $b\in B$ such that

$$P(a+b)>\frac{(|A|\,|B|)^{\frac{1}{2}}}{16\log N}. \tag{1}$$

Thus, in particular, if $|A|\gg N$ and $|B|\gg N$ then there exist $a\in A$ and $b\in B$ with

$$P(a+b)\gg\frac{N}{\log N}. \tag{2}$$

In part I of this paper [9], we obtained estimates for the greatest prime factor of sums of integers taken from k sets. In this paper we shall return to the case $k=2$. Our aim is to improve upon (1) when $|A|$ and $|B|$ are close to N. For example, we shall show that the right hand side of (2) may be replaced by N, which of course is best possible. Further we shall show that there exist many pairs (a, b) with a in A and b in B for which $P(a+b)$ is large.

Put

$$R=\frac{3N}{(|A|\,|B|)^{\frac{1}{2}}}.$$

Theorem. *Let N be a positive integer, let A and B be subsets of $\{1,\ldots,N\}$ and let ε be a positive real number. There exist effectively computable positive absolute constants c_1, c_2, c_3 and c_4 and a positive number N_1 which is effectively computable in terms of ε such that if $N>N_1$ and*

$$(|A|\,|B|)^{\frac{1}{2}}>N^{\frac{5}{6}+\varepsilon}, \tag{3}$$

*) The research of the second author was supported in part by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

then there exist at least $\frac{c_1|A|\,|B|}{\log N}$ *pairs* (a, b) *with* a *in* A *and* b *in* B *for which*

$$\frac{2c_2(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R} \geqq P(a+b) > \frac{c_2(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R}, \tag{4}$$

and there exist at least $\frac{c_3|A|\,|B|}{\log N}$ *pairs* (a_1, b_1) *with* a_1 *in* A *and* b_1 *in* B *for which*

$$\frac{2c_4(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R} \geqq P(a_1 - b_1) > \frac{c_4(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R}. \tag{5}$$

We remark that the estimates for the number of pairs (a, b) satisfying (4) and (5) can not be substantially improved. For example if $A = B = \{1, \ldots, N\}$ then the number of pairs satisfying (4) is at most $c_1^* \frac{N^2}{\log N} = c_1^* \frac{|A|\,|B|}{\log N}$, where c_1^* is a positive real number which is effectively computable in terms of c_2. Further, let T be a positive real number with

$$\frac{c_2(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R} > T > N^{\frac{5}{6}+\varepsilon}.$$

On applying the above theorem to subsets of A and B of the appropriate size we find that there exist a in A and b in B with $3T > P(a+b) > T$, provided that N is larger than a number which is effectively computable in terms of ε.

In particular, if (3) holds then for N sufficiently large there exist a in A and b in B such that

$$P(a+b) > \frac{c_2(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R}, \tag{6}$$

and there exist a_1 in A and b_1 in B with $a_1 \neq b_1$ such that

$$P(a_1 - b_1) > \frac{c_4(|A|\,|B|)^{\frac{1}{2}}}{\log R \log\log R}. \tag{7}$$

Thus if $|A| \gg N$ and $|B| \gg N$ then there exist a in A and b in B such that

$$P(a+b) \gg N,$$

and a_1 in A, b_1 in B with $a_1 \neq b_1$ such that

$$P(a_1 - b_1) \gg N.$$

Notice that (6) yields an improvement on (1) provided that

$$|A|\,|B| > N^2 \exp\left(-c_0\left(\frac{\log N}{\log\log N}\right)\right).$$

Furthermore, we remark that if A and B consist of all multiples of a positive integer t with $t \leqq N^{\frac{1}{2}}$ then for all a in A and b in B,

$$P(a \pm b) \leqq \max\left(P(t), 2\left[\frac{N}{t}\right]\right) \leqq 2\left[\frac{N}{t}\right] \leqq 2(|A|\,|B|)^{\frac{1}{2}},$$

hence (6) and (7) are nearly best possible even when $|A|\,|B| = o(N^2)$.

Perhaps for any $\varepsilon > 0$ there exist $N_0(\varepsilon)$ and $K = K(\varepsilon)$ such that if $N > N_0(\varepsilon)$ and $|A|\,|B| > KN$ then there exist $a \in A$ and $b \in B$ such that

$$P(a+b) > (2-\varepsilon)\,(|A|\,|B|)^{\frac{1}{2}}, \tag{8}$$

and $a_1 \in A$, $b_1 \in B$ with $a_1 \neq b_1$ such that

$$P(a_1 - b_1) > (1-\varepsilon)\,(|A|\,|B|)^{\frac{1}{2}}. \tag{9}$$

The following simple construction shows that the hypothesis $|A|\,|B| > KN$ is necessary in the above conjecture. Let γ be a real number with $0 < \gamma < \frac{1}{2}$ and let $n_1, \ldots, n_k$ be those positive integers n_i with $2 \leqq n_i \leqq N$ and $P(n_i) \leqq N^{\gamma}$. Put $A = \{1\}$ and $B = \{n_1 - 1, n_2 - 1, \ldots, n_k - 2\}$. By Lemma 3. 20 and Lemma 4. 7 of [7] there exists a positive number $c(\gamma)$ such that $|B| > c(\gamma) N$ for N sufficiently large. We then have

$$|A|\,|B| > c(\gamma)\, N,$$

while $P(a+b) \leqq N^{\gamma}$ for all $a \in A$ and $b \in B$.

§ 2. Preliminary lemmas

For any real number x let $[x]$ denote the greatest integer less than or equal to x, let $\{x\} = x - [x]$ denote the fractional part of x and let

$$\|x\| = \min(\{x\}, 1 - \{x\})$$

denote the distance from x to the nearest integer. Further denote $e^{2\pi i x}$ by $e(x)$.

Lemma 1. *Let X and Y be positive integers with $X < Y$. Then for any real number α we have*

$$\left| \sum_{X < n \leqq Y} e(n\alpha) \right| \leqq \min\left(Y - X, \frac{1}{2\|\alpha\|} \right).$$

Proof. See [8], p. 189.

Lemma 2. *Let V be a positive integer. Then for any real number α we have*

$$\left| \sum_{n=0}^{V-1} e(n\alpha) - V \right| \leqq 4V^2 |\alpha|.$$

Proof. See [1], Lemma 2.

For any positive integer n denote the number of integers less than or equal to n and coprime with n by $\phi(n)$. ϕ is Euler's phi function.

Lemma 3. *There exists an effectively computable positive real number c_5 such that*

$$\phi(n) > c_5 \frac{n}{\log\log n},$$

for $n \geqq 3$.

Proof. See [8], p. 24.

For any positive integer n, denote the number of positive integers which divide n by $\tau(n)$.

Lemma 4. *Let q be a positive integer and let u and v be real numbers with $v > 0$. Then*

$$\left| \sum_{\substack{u < k \leqq u+v \\ (k,q)=1}} 1 - v\frac{\phi(q)}{q} \right| \leqq \tau(q). \tag{10}$$

Proof.

$$\left| \sum_{\substack{u < k \leqq u+v \\ (k,q)=1}} 1 - v\frac{\phi(q)}{q} \right| = \left| \sum_{u < k \leqq u+v} \sum_{d \mid (k,q)} \mu(d) - v\frac{\phi(q)}{q} \right|$$

$$= \left| \sum_{d \mid q} \mu(d) \sum_{\substack{u < k \leqq u+v \\ d \mid k}} 1 - v \sum_{d \mid q} \frac{\mu(d)}{d} \right|$$

$$= \left| \sum_{d \mid q} \mu(d) \left(\sum_{\substack{u < k \leqq u+v \\ d \mid k}} 1 - \frac{v}{d} \right) \right| \leqq \sum_{d \mid q} 1 = \tau(q).$$

Lemma 5. *There exists an effectively computable positive real number c_6 such that for any integers a and b with $b \geqq 2$,*

$$\frac{b}{\phi(b)} \sum_{\substack{1 \leqq n \leqq b \\ (n+a,b)=1}} \frac{1}{n} < c_6 \log b.$$

Proof. Since the result plainly holds for $b = 2$ we may assume that $b \geqq 3$. First we notice that the result holds with $a = 0$. We have

$$\sum_{\substack{n=1 \\ (n,b)=1}}^{b} \frac{1}{n} < \prod_{\substack{p \leqq b \\ p \nmid b}} \left(1 - \frac{1}{p}\right)^{-1}$$

hence

$$\frac{b}{\phi(b)} \sum_{\substack{n=1 \\ (n,b)=1}}^{b} \frac{1}{n} < \prod_{p \leqq b} \left(1 - \frac{1}{p}\right)^{-1}$$

and, by Theorem 429 of [4],

$$\frac{b}{\phi(b)} \sum_{\substack{n=1 \\ (n,b)=1}}^{n} \frac{1}{n} < c_7 \log b. \tag{11}$$

Next put

$$t = \max_{u,v \in \mathbb{Z}^+} \left| \sum_{\substack{u < k \leqq u+v \\ (k,b)=1}} 1 - \sum_{\substack{0 < k \leqq v \\ (k,b)=1}} 1 \right|,$$

where $\mathbb{Z}^+$ denotes the set of positive integers. We have

$$\sum_{\substack{n=1\\(n+a,\,b)=1}}^{b}\frac{1}{n}\leqq\sum_{\substack{n=1\\(n,\,b)=1}}^{b}\frac{1}{n}+\sum_{n=1}^{t}\frac{1}{n},$$

since the j-th term in the sum on the left above is at most $\frac{1}{j}$ for $1\leqq j\leqq t$ and is at most the $(j-t)$-th term in $\sum_{\substack{n=1\\(n,\,b)=1}}^{b}\frac{1}{n}$ for $t<j\leqq\phi(b)$. Therefore, by (11),

$$\frac{b}{\phi(b)}\sum_{\substack{n=1\\(n+a,\,b)=1}}^{b}\frac{1}{n}<c_7\log b+\frac{b}{\phi(b)}(1+\log t)$$

which, by Lemma 3, is

$$<c_7\log b+c_8\log\log b(1+\log t).$$

It follows from Lemma 4 that $t\leqq 2\tau(b)$. Further, by Theorem 317 of [4],

$$\tau(b)\leqq b^{\frac{c_9}{\log\log b}},$$

and thus $(1+\log t)\leqq c_{10}\frac{\log b}{\log\log b}$. Therefore

$$\frac{b}{\phi(b)}\sum_{\substack{n=1\\(n+a,\,b)=1}}^{b}\frac{1}{n}<c_{11}\log b,$$

as required.

Lemma 6. *Let h, a and q be integers with $a>0$, $q>1$ and $(a,q)=1$. Let $\rho(n)$ be a real valued function defined for those integers n with $h\leqq n<h+q$ and $(n,q)=1$. Put*

$$\lambda=\max_{\substack{h\leqq n<h+q\\(n,\,q)=1}}\rho(n)-\min_{\substack{h\leqq n<h+q\\(n,\,q)=1}}\rho(n),$$

and

$$\psi(n)=\frac{1}{q}(an+\rho(n)).$$

There is an effectively computable positive absolute constant c_{12} such that if $\lambda\leqq 1$ and if E is a real number satisfying $2\leqq E\leqq q$ then

$$\sum_{\substack{h\leqq n<h+q\\(n,\,q)=1}}\min\left(E,\frac{1}{\|\psi(n)\|}\right)<c_{12}\phi(q)\log E.$$

Proof. Put $r=\left[\min_{\substack{h\leqq n<h+q\\(n,\,q)=1}}\rho(n)\right]$ and $\rho_1(n)=\rho(n)-r$. Note

$$0\leqq\rho_1(n)\leqq\lambda+1\leqq 2.$$

We have $\psi(n)=\frac{1}{q}\left((an+r)+\rho_1(n)\right)$ and so

$$\frac{an+r}{q}\leqq\psi(n)\leqq\frac{an+r+2}{q},$$

hence

$$\frac{1}{\|\psi(n)\|}\leqq\max\left(\frac{1}{\left\|\frac{an+r}{q}\right\|},\frac{1}{\left\|\frac{an+r+1}{q}\right\|},\frac{1}{\left\|\frac{an+r+2}{q}\right\|}\right),$$

subject to the convention that $a\leqq\max\left(\frac{1}{0},b\right)$ and $\frac{1}{0}\leqq\max\left(\frac{1}{0},a\right)$ for all real numbers a and b. Thus

$$\sum_{\substack{h\leqq n<h+q\\(n,q)=1}}\min\left(E,\frac{1}{\|\psi(n)\|}\right)\leqq\sum_{\substack{h\leqq n<h+q\\(n,q)=1}}\sum_{i=0}^{2}\min\left(E,\left\|\frac{an+r+i}{q}\right\|^{-1}\right)$$

$$\leqq 3\max_{j\in Z}\sum_{\substack{h\leqq n<h+q\\(n,q)=1}}\min\left(E,\left\|\frac{an+j}{q}\right\|^{-1}\right)$$

which, since $(a,q)=1$, is

$$\leqq 3\max_{j\in Z}\sum_{\substack{-j\leqq v<-j+q\\(v,q)=1}}\min\left(E,\left\|\frac{v+j}{q}\right\|^{-1}\right)$$

$$\leqq 3\max_{j\in Z}\left(E+\sum_{\substack{1\leqq t\leqq q-1\\(t-j,q)=1}}\min\left(E,\left\|\frac{t}{q}\right\|^{-1}\right)\right),$$

and since $\left\|\frac{t}{q}\right\|^{-1}\leqq\max\left(\frac{q}{t},\frac{q}{q-t}\right)$ for $1\leqq t\leqq q-1$, we have, on putting $q-t=x$, that the above is

$$\leqq 3\max_{j\in Z}\left(E+\sum_{\substack{1\leqq t\leqq q-1\\(t-j,q)=1}}\min\left(E,\frac{q}{t}\right)+\sum_{\substack{1\leqq x\leqq q-1\\(x+j,q)=1}}\min\left(E,\frac{q}{x}\right)\right)$$

$$\leqq 6\max_{j\in Z}\left(E+\sum_{\substack{1\leqq t\leqq q-1\\(t-j,q)=1}}\min\left(E,\frac{q}{t}\right)\right),\tag{12}$$

$$\leqq 6E+6q\max_{j\in Z}\left(\sum_{\substack{1\leqq t\leqq q-1\\(t-j,q)=1}}\frac{1}{t}\right),$$

which, by Lemma 5 is

$$\leqq 6E+6c_6\phi(q)\log q.\tag{13}$$

If $q\geqq E\geqq q^{\frac{1}{3}}$ then it follows from Lemma 3 and (13) that

$$\sum_{\substack{h\leqq n<h+q\\(n,q)=1}}\min\left(E,\frac{1}{\|\psi(n)\|}\right)\leqq c_{13}\phi(q)\log E.\tag{14}$$

On the other hand if $2 \leqq E \leqq q^{\frac{1}{3}}$ then the right hand side of inequality (12) is

$$\leqq 6 \max_{j \in Z} \left(E + \sum_{\substack{1 \leqq t < \frac{q}{E} \\ (t-j,q)=1}} E + \sum_{m=1}^{[E]} \sum_{\substack{\frac{mq}{E} \leqq t < (m+1)\frac{q}{E} \\ (t-j,q)=1}} \frac{q}{t} \right)$$

which, by Lemma 4, is

$$\leqq 6 \left(E + E\left(\frac{\phi(q)}{E} + \tau(q) \right) + \sum_{m=1}^{[E]} \frac{E}{m} \left(\frac{\phi(q)}{E} + \tau(q) \right) \right).$$

Since $\tau(q) \leqq 2q^{\frac{1}{2}}$ for all positive integers q, and $E < q^{\frac{1}{3}}$ it follows from Lemma 3 that

$$\frac{\phi(q)}{E} + \tau(q) < c_{14} \frac{\phi(q)}{E}.$$

Therefore

$$\sum_{\substack{h \leqq n < h+q \\ (n,q)=1}} \min\left(E, \frac{1}{\|\psi(n)\|} \right) \leqq 6 \left(E + c_{14}\phi(q) + c_{14}\phi(q) \sum_{m=1}^{[E]} \frac{1}{m} \right)$$

which by Lemma 3 is

$$\leqq c_{15}\phi(q) + c_{16}\phi(q) \log E \leqq c_{17}\phi(q) \log E. \tag{15}$$

Our result follows from (14) and (15).

We shall also require the Brun-Titchmarsh theorem, a result of Heath-Brown and Iwaniec and a refinement, due to Vaughan, of Vinogradov's fundamental lemma on exponential sums.

Let x be a positive real number and let l and k be positive integers. As usual we denote the number of primes less than or equal to x by $\pi(x)$ and the number of primes less than or equal to x and congruent to l modulo k by $\pi(x, k, l)$.

Lemma 7 (Brun-Titchmarsh Theorem). *Let x and y be positive real numbers and let k and l be relatively prime positive integers with $y > k$. Then*

$$\pi(x+y, k, l) - \pi(x, k, l) < \frac{2y}{\phi(k) \log\left(\frac{y}{k}\right)}.$$

Proof. See Theorem 2 of [6].

Lemma 8. *Given $\varepsilon > 0$ there exist positive real numbers $C_0 = C_0(\varepsilon)$ and $x_0 = x_0(\varepsilon)$, which are effectively computable in terms of ε, such that if $y \geqq x^{\frac{11}{20}+\varepsilon}$ and $x > x_0$ then*

$$\pi(x+y) - \pi(x) > C_0 \frac{y}{\log y}.$$

Proof. See [5].

We remark that for our purposes we require Lemma 8 only for the range $y \geqq x^{\frac{3}{5}}$.

Lemma 9. *If α is a real number and a, q and N are positive integers with $(a, q)=1$, $q \leqq N$ and $\left|\alpha-\frac{a}{q}\right| \leqq q^{-2}$ then*

$$\left|\sum_{p \leqq N} e(p\alpha)\right| < c_{18}(\log N)^4 (Nq^{-\frac{1}{2}} + N^{\frac{4}{5}} + N^{\frac{1}{2}} q^{\frac{1}{2}}),$$

where c_{18} is an effectively computable positive absolute constant; the summation above is over primes p with $p \leqq N$.

Proof. This follows from Theorem 3.1 of [10] by partial summation.

§ 3. Further preliminaries

In order to prove our main theorem we shall employ the Hardy-Littlewood method much as in [2]. In fact, apart from the values of the parameters, we shall start out from the same integral. However the integral must be estimated in a much more elaborate way. In particular, the treatment of the „major arcs" requires new ideas.

Let ε be a positive real number less than $\frac{1}{6}$ and let $N_1, N_2, \ldots$ denote positive numbers which are effectively computable in terms of ε. Put

$$y = \omega R \log R \log\log R$$

where ω is a positive real number larger than 400 which is effectively computable. Since $R \geqq 3$, $y \geqq (3 \log 3 \log\log 3)\,\omega \geqq \frac{\omega}{4}$, hence

$$y > 100. \tag{16}$$

Further, if (3) holds and $N > N_5$ *then*

$$y < 3\omega N^{\frac{1}{6}-\varepsilon} \log N^{\frac{1}{6}} \log\log N^{\frac{1}{6}} < \frac{N^{\frac{1}{6}-\frac{\varepsilon}{2}}}{2(\log N)^5} \tag{17}$$

and so

$$y < N^{\frac{1}{6}}. \tag{18}$$

We shall first establish (4). To do so it suffices to show that there exist at least $c_1 \frac{|A|\,|B|}{\log N}$ pairs (a, b) with a in A and b in B for which

$$\frac{4N}{y} \geqq P(a+b) > \frac{2N}{y}. \tag{19}$$

To this end we introduce the following notation. Put

$$Q = \frac{N}{y^3 (\log N)^{10}}, \quad \delta = \frac{y}{8N}, \quad U = \left[\frac{N}{y^2}\right]$$

and, for each positive integer n,

$$d_n = \begin{cases} 1 & \text{if } n = mp \text{ with } 1 \leqq m \leqq y \text{ and } p \text{ a prime such that } \frac{2N}{y} < p \leqq \frac{4N}{y}, \\ 0 & \text{otherwise.} \end{cases}$$

Next put

$$S(\alpha)=\sum_{n=1}^{4N} d_n e(n\alpha),$$

$$S=S(0)=\sum_{n=1}^{4N} d_n,$$

$$U(\alpha)=\sum_{n=0}^{U-1} e(n\alpha),$$

and, since $d_n=0$ if $n<1$ or $n>4N$, write

$$S(\alpha)\,U(\alpha)=\sum_{n=1}^{4N+U-1} v_n e(n\alpha) \quad \text{where} \quad v_n=\sum_{j=n-U+1}^{n} d_j.$$

Further, put

$$F(\alpha)=\sum_{a\in A} e(a\alpha), \quad G(\alpha)=\sum_{b\in B} e(b\alpha)$$

and

$$H(\alpha)=F(\alpha)\,G(\alpha)=\sum_{a\in A,\,b\in B} e\big((a+b)\,\alpha\big)=\sum_{n=1}^{2N} h_n e(n\alpha)$$

where

$$h_n=\sum_{\substack{a+b=n\\ a\in A,\,b\in B}} 1.$$

Finally, define J by

$$J=\int_0^1 F(\alpha)\,G(\alpha)\,S(-\alpha)\,d\alpha.$$

Observe that

$$J=\int_0^1 H(\alpha)\,S(-\alpha)\,d\alpha=\int_0^1 \sum_{n=1}^{2N}\sum_{m=1}^{4N} h_n d_m e\big((n-m)\,\alpha\big)\,d\alpha$$

$$=\sum_{n=1}^{2N} h_n d_n.$$

Also note that $d_n>0$ implies that $\dfrac{4N}{y}\geqq P(n)>\dfrac{2N}{y}$ while $h_n>0$ implies that n can be expressed as $a+b$ for some $a\in A$ and $b\in B$. Thus, to establish (19) and hence also (4), it suffices to show that

$$J>c_1\frac{|A|\,|B|}{\log N}. \tag{20}$$

In order to prove (20) we shall first establish estimates for S, $S(\alpha)$ and v_n.

It follows from (18) that, for $N>N_5$, $y<\dfrac{2N}{y}$ and therefore

$$S(\alpha)=\sum_{m\leqq y}\sum_{\frac{2N}{y}<p\leqq\frac{4N}{y}} e(mp\alpha). \tag{21}$$

Lemma 10. For $N > N_5$ we have

$$S < 10\frac{N}{\log N}.$$

Proof. By (21), for $N > N_5$,

$$S = \sum_{n=1}^{4N} d_n = \left(\sum_{1 \leqq m \leqq y} 1\right)\left(\sum_{\frac{2N}{y} < p \leqq \frac{4N}{y}} 1\right) \leqq y\pi\left(\frac{4N}{y}\right),$$

which, by Lemma 7 with $k=1$, is

$$\leqq \frac{8N}{\log\left(\frac{4N}{y}\right)} < \frac{8N}{\log(N/N^{\frac{1}{6}})} < 10\frac{N}{\log N}.$$

Lemma 11. *If $N > N_6$ then*

$$|S(\alpha)| < c_{19}\frac{N \log y \log\log y}{y \log N},$$

for $\delta < \alpha < 1-\delta$.

Proof. Let T_1 denote the set of those α in the interval $(\delta, 1-\delta)$ for which for all integers n with $1 \leqq n \leqq y$ there exist positive integers r_n and s_n with $(r_n, s_n) = 1$,

$$\left|n\alpha - \frac{r_n}{s_n}\right| < \frac{1}{s_n^2}, \tag{22}$$

and

$$y^2(\log N)^{10} \leqq s_n \leqq Q. \tag{23}$$

Put $T' = (\delta, 1-\delta) - T_1$, so that T' consists of the real numbers α in $(\delta, 1-\delta)$ which are not in T_1. If $\alpha \in T'$, then for some integer n^* with $1 \leqq n^* \leqq y$ there exist no coprime positive integers r_{n^*}, s_{n^*} satisfying (22) and (23) with n^* in place of n. By Dirichlet's theorem there exist integers u and v with

$$\left|n^*\alpha - \frac{u}{v}\right| < \frac{1}{vQ}, \tag{24}$$

$0 \leqq u$, $0 < v \leqq Q$ and $(u, v) = 1$. Note that

$$\left|n^*\alpha - \frac{u}{v}\right| < \frac{1}{v^2},$$

and therefore that $v < y^2(\log N)^{10}$. It follows directly from (24) that

$$\left|\alpha - \frac{u}{n^*v}\right| < \frac{1}{n^*vQ}$$

hence, on writing $\frac{u}{n^* v}$ in the form $\frac{a}{b}$ with a and b coprime, $a \geqq 0$ and $b > 0$ we see that

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{bQ} \tag{25}$$

with

$$b \leqq n^* v \leqq y^3 (\log N)^{10}. \tag{26}$$

To each α in T' we shall associate a pair of coprime integers a and b with $a \geqq 0$ and $b > 0$ satisfying (25) and (26) and we shall put

$$\beta = \alpha - \frac{a}{b}.$$

Let us define subsets T_2, T_3 and T_4 of T' in the following way:

$$T_2 = \left\{\alpha \in T' \mid 1 \leqq b \leqq y,\ |\beta| \leqq \frac{y}{8bN}\right\}.$$

$$T_3 = \left\{\alpha \in T' \mid 1 \leqq b \leqq y,\ |\beta| > \frac{y}{8bN}\right\}.$$

$$T_4 = \{\alpha \in T' \mid y < b\}.$$

Since $(\delta, 1-\delta) = T_1 \cup T_2 \cup T_3 \cup T_4$ it suffices to show that

$$\max_{\alpha \in T_i} |S(\alpha)| < c_{19} \frac{N \log y \log\log y}{y \log N}, \tag{27}$$

for $i = 1, 2, 3, 4$ and for $N > N_6$.

We shall first establish (27) for $i = 1$. Accordingly assume that $\alpha \in T_1$. By (21), for $N > N_5$,

$$|S(\alpha)| = \left|\sum_{n \leqq y} \sum_{\frac{2N}{y} < p \leqq \frac{4N}{y}} e(np\alpha)\right| \tag{28}$$

$$\leqq \sum_{n \leqq y} \left(\left|\sum_{p \leqq \frac{4N}{y}} e\big(p(n\alpha)\big)\right| + \left|\sum_{p \leqq \frac{2N}{y}} e\big(p(n\alpha)\big)\right|\right),$$

which by Lemma 9, (22) and (23), is

$$\leqq \sum_{n \leqq y} c_{20} \left(\log \frac{4N}{y}\right)^4 \left(\frac{4N}{y}\left(y^2 (\log N)^{10}\right)^{-\frac{1}{2}} + \left(\frac{4N}{y}\right)^{\frac{4}{5}} + \left(\frac{4N}{y}\right)^{\frac{1}{2}} Q^{\frac{1}{2}}\right)$$

$$\leqq c_{21}\, y (\log N)^4 \left(\frac{6N}{y^2 (\log N)^5} + \left(\frac{4N}{y}\right)^{\frac{4}{5}}\right)$$

and, by (17),

$$\leqq c_{22} \frac{N}{y \log N}, \tag{29}$$

for $N > N_7$.

Now assume that $\alpha \in T_2$. Notice that we may assume that $b > 1$ since if $b = 1$ then $|\beta| \leqq \frac{y}{8N} = \delta$ and consequently α is not in $(\delta, 1-\delta)$. Further since $b \neq 1$ we may assume that $a \neq 0$. We have, from (28),

$$|S(\alpha)| \leqq \sum_{\frac{2N}{y} < p \leqq \frac{4N}{y}} \left| \sum_{n \leqq y} e(np\alpha) \right|$$

which, by Lemma 1, is

$$\leqq \sum_{\frac{2N}{y} < p \leqq \frac{4N}{y}} \min\left(y, \frac{1}{2\|p\alpha\|}\right).$$

It follows from (18) and (26) that if p is a prime with $p > \frac{2N}{y}$ and $N > N_8$ then $(p, b) = 1$. Thus for $N > N_9$

$$|S(\alpha)| \leqq \sum_{\substack{p \leqq \frac{4N}{y} \\ (p,b)=1}} \min\left(y, \frac{1}{2\|p\alpha\|}\right). \tag{30}$$

Notice that

$$\|p\alpha\| = \left\| p\left(\frac{a}{b} + \beta\right) \right\| \geqq \left\| \frac{ap}{b} \right\| - p|\beta| \geqq \left\| \frac{ap}{b} \right\| - \left(\frac{4N}{y}\right)\left(\frac{y}{8bN}\right)$$
$$= \left\| \frac{ap}{b} \right\| - \frac{1}{2b} \geqq \frac{1}{2} \left\| \frac{ap}{b} \right\|,$$

since $b > 1$ and $(ap, b) = 1$. Thus, from (30), for $N > N_{10}$

$$|S(\alpha)| \leqq \sum_{\substack{p \leqq \frac{4N}{y} \\ (p,b)=1}} \min\left(y, \frac{1}{\left\| \frac{ap}{b} \right\|}\right)$$

$$\leqq \sum_{\substack{0 \leqq h < b \\ (h,b)=1}} \left(\sum_{\substack{p \leqq \frac{4N}{y} \\ ap \equiv h \pmod b}} \frac{1}{\left\| \frac{h}{b} \right\|} \right)$$

$$\leqq \left(\max_{\substack{0 < l < b \\ (l,b)=1}} \pi\left(\frac{4N}{y}, b, l\right) \right) \sum_{\substack{0 \leqq h < b \\ (h,b)=1}} \frac{1}{\left\| \frac{h}{b} \right\|}$$

which, by Lemma 7, (18) and (26), is

$$\leqq \frac{8N}{y\phi(b) \log\left(\frac{4N}{by}\right)} \cdot 2 \sum_{\substack{1 \leqq h \leqq \frac{b}{2} \\ (h,b)=1}} \frac{b}{h}$$

and, by Lemma 5, this is

$$\leqq c_{23} \frac{N}{y \log N} \log b \leqq c_{23} \frac{N}{y \log N} \log y, \tag{31}$$

as required.

We shall assume next that $\alpha \in T_3$ hence

$$\frac{y}{8bN} < |\beta| < \frac{1}{bQ}. \tag{32}$$

Put

$$L = \frac{1}{2b|\beta|}. \tag{33}$$

It follows from (32) that

$$\frac{Q}{2} < L < \frac{4N}{y}.$$

Then, from (30), for $N > N_9$,

$$\begin{aligned} |S(\alpha)| &\leqq \sum_{\substack{p \leqq \frac{4N}{y} \\ (p,b)=1}} \min\left(y, \frac{1}{2\|p\alpha\|}\right) \\ &\leqq \sum_{j=1}^{\left[\frac{4N}{Ly}\right]+1} \sum_{\substack{(j-1)L < p \leqq jL \\ (p,b)=1}} \min\left(y, \frac{1}{2\|p\alpha\|}\right) \\ &= \sum_{j=1}^{\left[\frac{4N}{Ly}\right]+1} \sum_{k=1}^{2y} \sum_{\substack{(j-1)L < p \leqq jL \\ (p,b)=1 \\ \frac{k-1}{2y} \leqq \{p\alpha\} < \frac{k}{2y}}} \min\left(y, \frac{1}{2\|p\alpha\|}\right). \end{aligned}$$

Since $\frac{k-1}{2y} \leqq \{p\alpha\} < \frac{k}{2y}$ implies that

$$\frac{1}{\|p\alpha\|} \leqq \frac{1}{\left\|\frac{k-1}{2y}\right\|} + \frac{1}{\left\|\frac{k}{2y}\right\|},$$

where as before we write $a \leqq \frac{1}{0} + b$ and $\frac{1}{0} \leqq \frac{1}{0} + a$ for all real numbers a and b, we have

$$|S(\alpha)| \leqq \sum_{j=1}^{\left[\frac{4N}{Ly}\right]+1} \sum_{k=1}^{2y} \left(\min\left(y, \frac{1}{2\left\|\frac{k-1}{2y}\right\|}\right) + \min\left(y, \frac{1}{2\left\|\frac{k}{2y}\right\|}\right)\right) \sum_{\substack{(j-1)L < p \leqq jL \\ (p,b)=1 \\ \frac{k-1}{2y} \leqq \{p\alpha\} < \frac{k}{2y}}} 1. \tag{34}$$

If p and p_0 are primes with $(j-1)L < p \leqq jL$, $\frac{k-1}{2y} \leqq \{p\alpha\} < \frac{k}{2y}$ and $(j-1)L < p_0 \leqq jL$, $\frac{k-1}{2y} \leqq \{p_0\alpha\} < \frac{k}{2y}$ then

$$\begin{aligned} \frac{1}{2y} > \|(p-p_0)\alpha\| &= \left\|(p-p_0)\left(\frac{a}{b}+\beta\right)\right\| \geqq \left\|(p-p_0)\frac{a}{b}\right\| - |p-p_0|\,|\beta| \\ &> \left\|(p-p_0)\frac{a}{b}\right\| - L|\beta| = \left\|(p-p_0)\frac{a}{b}\right\| - \frac{1}{2b}. \end{aligned}$$

Thus

$$\left\|(p-p_0)\frac{a}{b}\right\| < \frac{1}{2y}+\frac{1}{2b} \leqq \frac{1}{b},$$

whence

$$p \equiv p_0 \pmod b. \tag{35}$$

Therefore

$$\frac{1}{2y} > \|p\alpha - p_0\alpha\| = \left\|(p-p_0)\frac{a}{b}+(p-p_0)\beta\right\| = \|(p-p_0)\beta\|. \tag{36}$$

Since

$$|(p-p_0)\beta| < L|\beta| = \frac{1}{2b} \leqq \frac{1}{2},$$

it follows from (36) that

$$\frac{1}{2y} > |p-p_0|\,|\beta|,$$

hence

$$|p-p_0| < \frac{1}{2|\beta|y}.$$

Thus, either there are no primes p with $(j-1)L < p \leqq jL$, $(p, b)=1$ and $\frac{k-1}{2y} \leqq \{p\alpha\} < \frac{k}{2y}$, or for some p_0 we have

$$\begin{aligned} \sum_{\substack{(j-1)L<p\leqq jL \\ (p,b)=1 \\ \frac{k-1}{2y}\leqq\{p\alpha\}<\frac{k}{2y}}} 1 &\leqq \sum_{\substack{p_0-\frac{1}{2|\beta|y}<p<p_0+\frac{1}{2|\beta|y} \\ p\equiv p_0 \pmod b}} 1 \\ &\leqq \pi\left(p_0+\frac{1}{2|\beta|y}, b, p_0\right)-\pi\left(p_0-\frac{1}{2|\beta|y}, b, p_0\right). \end{aligned} \tag{37}$$

By (32), $\frac{1}{|\beta|y} > \frac{bQ}{y}$. Thus, for $N > N_{11}$, the right hand side of inequality (37) is, by (18) and Lemma 7,

$$\leqq \frac{\frac{2}{|\beta|y}}{\phi(b)\log\left(\frac{1}{|\beta|yb}\right)}$$

and, by (33), is

$$\leqq \frac{4bL}{y\phi(b)\log\left(\frac{Q}{y}\right)} \leqq \frac{c_{24}bL}{y\phi(b)\log N}.$$

It now follows from (34), that

$$|S(\alpha)| \leqq \sum_{j=1}^{\left[\frac{4N}{Ly}\right]+1} \sum_{k=1}^{2y} \left(\min\left(y, \frac{1}{2\left\|\frac{k-1}{2y}\right\|}\right) + \min\left(y, \frac{1}{2\left\|\frac{k}{2y}\right\|}\right)\right) \frac{c_{24} bL}{y\phi(b)\log N}$$

$$\leqq \left(\left[\frac{4N}{Ly}\right]+1\right) \frac{c_{25} bL}{y\phi(b)\log N} \sum_{k=0}^{y} \min\left(y, \frac{1}{2\left\|\frac{k}{2y}\right\|}\right)$$

$$\leqq c_{26} \frac{Nb}{y^2\phi(b)\log N}\left(y + \sum_{k=1}^{y} \frac{y}{k}\right) \leqq c_{27} \frac{N\log yb}{y\log N\phi(b)}$$

which, by Lemma 3, is

$$\leqq c_{28} \frac{N\log y \log\log b}{y\log N}.$$

Since $b \leqq y$ we have

$$|S(\alpha)| \leqq \frac{c_{28} N\log y\log\log y}{y\log N}. \tag{38}$$

for $\alpha \in T_3$ provided that $N > N_{12}$.

Finally we assume that $\alpha \in T_4$. Put

$$L_1 = \min\left(\frac{N}{y}, \frac{1}{2|\beta| y}\right).$$

Then, by (30), for $N > N_9$,

$$|S(\alpha)| \leqq \sum_{\substack{p \leqq \frac{4N}{y} \\ (p,b)=1}} \min\left(y, \frac{1}{2p\|\alpha\|}\right) \tag{39}$$

$$\leqq \sum_{j=1}^{\left[\frac{4N}{L_1 y}\right]+1} \sum_{\substack{(j-1)L_1 < p \leqq jL_1 \\ (p,b)=1}} \min\left(y, \frac{1}{2\|p\alpha\|}\right).$$

Now if $\frac{1}{2\|p\alpha\|} < y$ with $(j-1)L_1 < p \leqq jL_1$ and n is defined by $p \equiv n \pmod b$ with $jL_1 - b < n \leqq jL_1$ then

$$\|p\alpha\| = \left\|p\left(\frac{a}{b}+\beta\right)\right\| = \left\|\frac{an}{b} + n\beta + (p-n)\beta\right\|$$

$$\geqq \left\|\frac{1}{b}(an+nb\beta)\right\| - |p-n|\,|\beta|$$

and since $|p-n|\,|\beta| \leqq L_1|\beta| \leqq \frac{1}{2y} < \|p\alpha\|$ we have

$$2\|p\alpha\| \geqq \left\| \frac{1}{b}(an+nb\beta) \right\|,$$

hence

$$\min\left(y, \frac{1}{2\|p\alpha\|}\right) \leqq \min\left(y, \frac{1}{\left\|\frac{1}{b}(an+nb\beta)\right\|}\right).$$

Therefore, by (39),

$$|S(\alpha)| \leqq \sum_{j=1}^{\left[\frac{4N}{L_1 y}\right]+1} \sum_{\substack{jL_1-b<n\leqq jL_1 \\ (n,b)=1}} \min\left(y, \frac{1}{\left\|\frac{1}{b}(an+nb\beta)\right\|}\right) \sum_{\substack{(j-1)L_1<p\leqq jL_1 \\ p\equiv n \pmod b}} 1. \tag{40}$$

We see from (25) and the fact that $b>y$ that $\frac{1}{2|\beta|\,y} > \frac{Q}{2}$ hence that $L_1 > \frac{Q}{2}$. For $N > N_5$ we have, by (17) and (26),

$$\frac{L_1}{b} > \frac{Q}{2b} \geqq \frac{N}{2y^6(\log N)^{20}} \geqq N^{3\varepsilon}, \tag{41}$$

whence, from Lemma 7,

$$\sum_{\substack{(j-1)L_1<p\leqq jL_1 \\ p\equiv n \pmod b}} 1 < \frac{2L_1}{\phi(b)\log\left(\frac{L_1}{b}\right)} \leqq \frac{L_1}{\varepsilon\phi(b)\log N}. \tag{42}$$

Combining (40) and (42) we obtain

$$|S(\alpha)| \leqq \frac{L_1}{\varepsilon\phi(b)\log N} \sum_{j=1}^{\left[\frac{4N}{L_1 y}\right]+1} \sum_{\substack{jL_1-b<n\leqq jL_1 \\ (n,b)=1}} \min\left(y, \frac{1}{\left\|\frac{1}{b}(an+nb\beta)\right\|}\right).$$

We may estimate the inner sum above by means of Lemma 6 with $h = jL_1 - b + 1$, $q = b$ and $\rho(n) = nb\beta$. Then, by (17) and (26),

$$\lambda = \max_{jL_1-b<n\leqq jL_1} nb\beta - \min_{jL_1-b<n\leqq jL_1} nb\beta$$

$$\leqq b^2|\beta| < \frac{b}{Q} < \frac{y^6(\log N)^{20}}{N} < 1,$$

for $N > N_5$. Thus

$$|S(\alpha)| \leqq \frac{L_1}{\varepsilon\phi(b)\log N}\left(\left[\frac{4N}{L_1 y}\right]+1\right) c_{12}\phi(b)\log y,$$

and, since $L_1 \leqq \frac{N}{y}$,

$$|S(\alpha)| \leqq \frac{c_{29} N \log y}{\varepsilon y \log N}. \tag{43}$$

If $y<\frac{N^{\frac{1}{12}}}{2(\log N)^4}$ then we may replace 3ε in (41) by $\frac{1}{2}$ and consequently ε in (43) by 1. On the other hand if $y\geqq\frac{N^{\frac{1}{12}}}{2(\log N)^4}$ then certainly $\frac{1}{\varepsilon}<\log\log y$ for $N>N_{13}$. Thus Lemma 11 holds for $\alpha\in T_4$ and, by (29), (31) and (38), the proof is complete.

Lemma 12. *If $N>N_{14}$ and n is an integer satisfying $\frac{30N}{y}<n\leqq 2N$ then*

$$v_n>c_{30}\frac{N}{y^2\log N}.$$

Proof. If n satisfies $\frac{30N}{y}<n\leqq 2N$ then for $N>N_{15}$,

$$v_n=\sum_{j=n-U+1}^{n}d_j=\sum_{\substack{n-U<mp\leqq n\\ m\leqq y\\ \frac{2N}{y}<p\leqq\frac{4N}{y}}}1=\sum_{m\leqq y}\sum_{\max\left(\frac{n-U}{m},\frac{2N}{y}\right)<p\leqq\min\left(\frac{n}{m},\frac{4N}{y}\right)}1.$$

Notice that if $m\leqq\frac{11}{30}\frac{ny}{N}$ then, by (16),

$$\frac{n-U}{m}\geqq\frac{30}{11}\frac{N}{y}-\frac{N}{y^2m}\geqq\frac{30}{11}\frac{N}{y}\left(1-\frac{11}{30y}\right)\geqq\frac{2N}{y}.$$

Further if $\frac{9}{30}\frac{ny}{N}<m$ then

$$\frac{n}{m}<\frac{30}{9}\frac{N}{y}<\frac{4N}{y}.$$

Since $\frac{11}{30}\frac{ny}{N}\leqq\frac{22}{30}y<y$ we conclude that

$$v_n>\sum_{\frac{9}{30}\frac{ny}{N}<m\leqq\frac{11}{30}\frac{ny}{N}}\ \sum_{\frac{n-U}{m}<p\leqq\frac{n}{m}}1=\sum_{\frac{9}{30}\frac{ny}{N}<m\leqq\frac{11}{30}\frac{ny}{N}}\left(\pi\left(\frac{n}{m}\right)-\pi\left(\frac{n-U}{m}\right)\right).$$

We may now apply Lemma 8 with $x=\frac{n-U}{m}$ and $y=\frac{U}{m}$ since for $N>N_{16}$ and $m\leqq\frac{11}{30}\frac{ny}{N}$ we have, by (18),

$$\left(\frac{n-U}{m}\right)^{\frac{12}{21}}<\left(\frac{n}{m}\right)^{\frac{4}{7}}=\frac{(n^4m^3)^{\frac{1}{7}}}{m}\leqq\frac{n}{m}\left(\frac{11}{30}\frac{y}{N}\right)^{\frac{3}{7}}\leqq\frac{2N}{m}\left(\frac{11}{30}\frac{y}{N}\right)^{\frac{3}{7}}<\frac{U}{m}.$$

Thus, for $N > N_{17}$,

$$v_n > \sum_{\frac{9}{30}\frac{ny}{N} < m \leqq \frac{11}{30}\frac{ny}{N}} c_{31} \frac{U}{m \log\left(\frac{U}{m}\right)} > \frac{c_{31} U}{\log U} \sum_{\frac{9}{30}\frac{ny}{N} < m \leqq \frac{11}{30}\frac{ny}{N}} \frac{1}{m}$$

and, by (18), is

$$> c_{31} \frac{N}{y^2 \log N} \left(\frac{30}{11}\frac{N}{ny}\right)\left(\frac{2}{30}\frac{ny}{N} - 1\right) > c_{32} \frac{N}{y^2 \log N}\left(1 - \frac{15N}{ny}\right),$$

which, since $n > \frac{30N}{y}$, is

$$> c_{33} \frac{N}{y^2 \log N},$$

as required.

§ 4. Proof of the theorem

We shall first prove (4). We have for $N > N_{18}$,

$$\left| J - \frac{1}{U}\int_0^1 F(\alpha)\, G(\alpha)\, U(-\alpha)\, S(-\alpha)\, d\alpha \right|$$

$$= \left| \int_{-\delta}^{\delta} F(\alpha)\, G(\alpha)\, S(-\alpha)\left(1 - \frac{U(-\alpha)}{U}\right) d\alpha + \int_{\delta}^{1-\delta} F(\alpha)\, G(\alpha)\, S(-\alpha)\left(1 - \frac{U(-\alpha)}{U}\right) d\alpha \right|$$

$$\leqq \int_{-\delta}^{\delta} |F(\alpha)|\, |G(\alpha)|\, |S(-\alpha)| \frac{|U - U(-\alpha)|}{U}\, d\alpha$$

$$+ \int_{\delta}^{1-\delta} |F(\alpha)|\, |G(\alpha)|\, |S(-\alpha)| \left(1 + \left|\frac{U(-\alpha)}{U}\right|\right) d\alpha$$

which, by Lemma 2, is

$$\leqq \int_{-\delta}^{\delta} |F(\alpha)|\, |G(\alpha)|\, S \frac{4U^2|\alpha|}{U}\, d\alpha + \int_{\delta}^{1-\delta} |F(\alpha)|\, |G(\alpha)| \left(\max_{\delta \leqq \beta \leqq 1-\delta} |S(\beta)|\right) 2 d\alpha$$

by Lemmas 10 and 11, is

$$\leqq \int_{-\delta}^{\delta} |F(\alpha)|\, |G(\alpha)| \frac{40N}{\log N}\, U\delta\, d\alpha + \int_{\delta}^{1-\delta} |F(\alpha)|\, |G(\alpha)|\, 2c_{19} \frac{N \log y \log\log y}{y \log N}\, d\alpha$$

$$\leqq \left(\frac{5N}{y \log N} + 2c_{19} \frac{N \log y \log\log y}{y \log N}\right) \int_0^1 |F(\alpha)|\, |G(\alpha)|\, d\alpha$$

and by Cauchy's inequality and Parseval's formula is

$$\leqq c_{34} \frac{N \log y \log\log y}{y \log N} \left(\left(\int_0^1 |F(\alpha)|^2\, d\alpha\right)\left(\int_0^1 |G(\alpha)|^2\, d\alpha\right)\right)^{\frac{1}{2}},$$

and thus is

$$\leqq c_{34} \frac{N \log y \log\log y}{y \log N} (|A|\, |B|)^{\frac{1}{2}}. \tag{44}$$

Furthermore,

$$I=\int_0^1 F(\alpha)\,G(\alpha)\,U(-\alpha)\,S(-\alpha)\,d\alpha=\int_0^1\left(\sum_{n=1}^{2N}h_n e(n\alpha)\right)\left(\sum_{n=1}^{4N+U-1}v_n e(-n\alpha)\right)d\alpha$$
$$=\sum_{n=1}^{2N}h_n v_n.$$

Since h_n and v_n are non-negative for $n=1,\dots,2N$,

$$I\geqq\sum_{\frac{30N}{y}<n\leqq 2N}h_n v_n,$$

and by Lemma 12,

$$I\geqq\frac{c_{30}N}{y^2\log N}\sum_{\frac{30N}{y}<n\leqq 2N}h_n=\frac{c_{30}N}{y^2\log N}\sum_{\substack{a\in A,\,b\in B\\ \frac{30N}{y}<a+b\leqq 2N}}1.$$

Observe that $\dfrac{30N}{y}\leqq\dfrac{3(|A|\,|B|)^{\frac12}}{10}\leqq\dfrac{3}{10}\max(|A|,|B|)$ and thus

$$\sum_{\substack{a\in A,\,b\in B\\ \frac{30N}{y}<a+b\leqq 2N}}1\geqq\frac{7}{10}|A|\,|B|.$$

Therefore

$$I\geqq\frac{c_{35}N|A|\,|B|}{y^2\log N}. \tag{45}$$

It follows from (44) and (45) that

$$|J|\geqq\frac{|I|}{U}-c_{34}\frac{N\log y\log\log y}{y\log N}(|A|\,|B|)^{\frac12}$$
$$\geqq\frac{c_{35}}{\log N}\left(|A|\,|B|\;-c_{36}\frac{N\log y\log\log y}{y}(|A|\,|B|)^{\frac12}\right).$$

Since $y=\omega R\log R\;\log\log R<\omega R^3$ we have

$$|J|\geqq\frac{c_{35}|A|\,|B|}{\log N}\left(1-c_{36}\frac{\log\omega R^3\log\log\omega R^3}{3\omega\log R\log\log R}\right)$$
$$\geqq\frac{c_{35}|A|\,|B|}{\log N}\left(1-c_{37}\frac{\log\omega\log\log\omega}{\omega}\right).$$

On choosing ω sufficiently large we find that

$$|J|\geqq\frac{c_{35}|A|\,|B|}{2\log N}.$$

Since J is non-negative, (20) holds and this completes the proof of (4).

The proof of (5) is essentially the same as that of (4). First observe that we may assume, without loss of generality, that $|A| \geqq |B|$. Next define z to be the smallest positive integer for which

$$|A \cap \{1, \ldots, z\}| \geqq \frac{|A|}{2}.$$

Put

$$A_1 = A \cap \left\{1, \ldots, z - \left[\frac{3|A|}{10}\right] - 1\right\},$$

$$A_2 = A \cap \left\{z + \left[\frac{3|A|}{10}\right] + 1, \ldots, N\right\},$$

$$B_1 = B \cap \{1, \ldots, z\},$$

$$B_2 = B \cap \{z+1, \ldots, N\}.$$

Note that the minimum of $|A_1|$ and $|A_2|$ is at least $\frac{|A|}{10}$ for $N > N_{19}$. Further the maximum of $|B_1|$ and $|B_2|$ is at least $\frac{|B|}{2}$.

We shall consider separately the case $|B_1| \geqq |B_2|$ and the case $|B_1| < |B_2|$. Assume first that $|B_1| \geqq |B_2|$. Then, since $\frac{30N}{y} \leqq \frac{3}{10}|A|$,

$$\sum_{\substack{a \in A,\, b \in B \\ \frac{30N}{y} < a-b \leqq N}} 1 \geqq |A_2|\,|B_1| \geqq \frac{|A|\,|B|}{20}. \tag{46}$$

We now replace $G(\alpha)$ by $G(-\alpha)$ in the above argument. Then

$$F(\alpha)\,G(-\alpha) = \sum_{a \in A,\, b \in B} e((a-b)\,\alpha) = \sum_{n=-N}^{N} h_n' e(n\alpha),$$

where

$$h_n' = \sum_{\substack{a \in A,\, b \in B \\ a-b=n}} 1.$$

Further we put $J' = \int_0^1 F(\alpha)\,G(-\alpha)\,S(-\alpha)\,d\alpha$ and $I' = \int_0^1 F(\alpha)\,G(-\alpha)\,U(-\alpha)\,S(-\alpha)\,d\alpha$. As before it suffices to show that $J' > \frac{c_3|A|\,|B|}{\log N}$. We have $I' = \sum_{n=1}^{N} h_n' v_n$ hence, by Lemma 12,

$$I' \geqq \frac{c_{38}N}{y^2 \log N} \sum_{\frac{30N}{y} < n \leqq N} h_n'.$$

Thus, by (46),

$$I' \geqq c_{39} \frac{N|A|\,|B|}{y^2 \log N}, \tag{47}$$

and the required lower bound for J' follows as above.

Finally we consider the case $|B_1| < |B_2|$. In this case

$$\sum_{\substack{a\in A,\, b\in B \\ \frac{30N}{y} < b-a \leqq N}} 1 \geqq |A_1|\,|B_2| \geqq \frac{|A|\,|B|}{20}. \tag{48}$$

We then put

$$F(-\alpha)\,G(\alpha) = \sum_{a\in A,\, b\in B} e((b-a)\,\alpha) = \sum_{n=-N}^{N} h_n'' e(n\alpha),$$

where

$$h_n'' = \sum_{\substack{a\in A,\, b\in B \\ b-a=n}} 1.$$

Further we put $J'' = \int_0^1 F(-\alpha)\,G(\alpha)\,S(-\alpha)\,d\alpha$ and $I'' = \int_0^1 F(-\alpha)\,G(\alpha)\,U(-\alpha)\,S(-\alpha)\,d\alpha$. Employing Lemma 12 and (48) we find that (47) holds with I' replaced by I''. We may now argue as above to show that $J'' > c_3 \dfrac{|A|\,|B|}{\log N}$. Since $P(b-a) = P(a-b)$ our result now follows.

Remark. By modifying the proof of Lemma 11 it is possible to show that there exists an effectively computable positive absolute constant c such that if N is a positive integer and y is a real number with $3 \leqq y \leqq N^{\frac{1}{6}}$ then

$$\sum_{p\leqq N} \min(y, \|p\alpha\|^{-1}) < c\,\frac{N \log y \log\log y}{\log N},$$

for all real numbers α with $N^{-1} \leqq \alpha \leqq 1 - N^{-1}$; the summation above is over primes p with $p \leqq N$. We shall establish this result in a subsequent paper.

References

[1] *A. Balog* and *A. Sárközy*, On sums of sequences of integers. I, Acta Arithmetica **44** (1984), 73—86.
[2] *A. Balog* and *A. Sárközy*, On sums of sequences of integers. II, Acta Math. Hung. **44** (1984), 169—179.
[3] *A. Balog* and *A. Sárközy*, On sums of sequences of integers. III, Acta Math. Hung. **44** (1984), 339—349.
[4] *G. H. Hardy* and *E. M. Wright*, An introduction to the theory of numbers, 5th ed., Oxford 1979.
[5] *D. R. Heath-Brown* and *H. Iwaniec*, On the difference between consecutive primes, Invent. Math. **55** (1979), 49—69.
[6] *H. L. Montgomery* and *R. C. Vaughan*, The large sieve, Mathematika **20** (1973), 119—134.
[7] *K. K. Norton*, Numbers with small prime factors, and the least k-th power non-residue, Memoirs of the American Mathematical Society **106** (1971).
[8] *K. Prachar*, Primzahlverteilung, Berlin-Heidelberg-New York 1957.
[9] *A. Sárközy* and *C. L. Stewart*, On divisors of sums of integers. I, Acta Math. Hung., to appear.
[10] *R. C. Vaughan*, The Hardy-Littlewood method, Cambridge 1981.

Mathematical Institute of the Hungarian Academy of Sciences,
Reáltanoda u. 13—16, Budapest, H-1053, Hungary

Department of Pure Mathematics, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1

Eingegangen 6. Juni 1985

Cohomology of arithmetic groups and congruences between systems of Hecke eigenvalues

By *Avner Ash**) at Columbus and *Glenn Stevens**) at Boston

The present paper is the beginning of an investigation into the congruence properties of systems of eigenvalues of Hecke operators acting on cohomology groups associated to automorphic forms over a reductive algebraic group G.

When $G = GL(2)$ this amounts to a study of the congruence properties of q-expansions of classical modular newforms of weight $\geqq 2$. This theory has been researched extensively and has found numerous applications (see for example [9], [16], [17], [22], [24], [28], [29], [31], [34]). The cohomological approach to the theory was initiated in 1968 by Shimura [33] (see Hida [17]). Other cohomological attacks on congruences can be found in [13] and [20]. In [20] forms over unit groups in quaternion algebras are studied as well.

In this paper we broaden the method in order to treat arithmetic subgroups of other reductive groups.

In section 1 we set up some general machinery for treating the cohomology of a group Γ (possibly with twisted coefficients) along with the action of a commutative Hecke algebra of double cosets on it. We consider pairs of groups and coefficient $\mathbf{Z}$-modules related in such a way that we obtain (1) a homomorphism ι between the Hecke algebras; (2) an ι-equivariant map between the cohomology groups mod l; and (3) a way of lifting systems of Hecke eigenvalues mod l back to the integral cohomology. The net result is a pair of integral cohomology eigenclasses, one per group, and a congruence mod l between the associated systems of Hecke eigenvalues.

Two remarks deserve special emphasis. First, despite the congruence between the eigenvalues associated to these eigenclasses, there may be no direct relationship between the eigenclasses themselves (compare the remark following proposition 1. 2. 3). Second, the cohomology groups under consideration may contain nontrivial torsion elements (cf. section 3). In the general situation, one of our eigenclasses may be a torsion class even if the other is not.

---

*) Research partially supported by grants from the National Science Foundation.

Note that our Hecke algebras are assumed to be commutative, a condition which is satisfied in many important examples.

In section 2 we prove two results for arithmetic groups Γ generalizing known results for $GL(2)$. The first of these (theorem 2. 2) states that for fixed Γ and l, the set of systems of Hecke eigenvalues modulo l occurring in $\oplus H^*(\Gamma; E)$ is finite, where the direct sum is over all finite dimensional rational representations E of the ambient group which can be "reduced" modulo l. This may be viewed as a generalization of a theorem of Serre and Tate [31] (see also Jochnowitz [19]) which states that the set of systems of Hecke eigenvalues modulo l arising from modular forms of all weights and level 1 is finite. Our second result (theorem 2. 4) states that, given Γ, l sufficiently large, and E as above, there exists a subgroup Γ_1 of finite index in Γ and a trivial Γ_1-module F such that every system of Hecke eigenvalues occurring in $H^N(\Gamma; E)$ may be found modulo l in $H^N(\Gamma_1; F)$ where N is the virtual cohomological dimension of Γ. This generalizes results of Serre [28] and Serre-Fontaine [29] which state that modular eigenforms of arbitrary weight are congruent modulo l to weight two forms. In a later paper we will show how to get similar results in dimensions other than N. We close section 2 with an examination of the special case $G = GL(n)$.

One expects that our methods, when applied to specific groups, will yield stronger theorems than the general ones proved in § 2. In the special case $G = GL(2)$, for example, our methods can be refined to obtain more explicit statements about congruences among modular forms. We are also able to prove congruences between the algebraic parts of special values of associated L-functions. We will report on this in an upcoming paper.

In section 3, we use our methods along with some input from automorphic representation theory and l-adic Galois representation theory to prove (theorem 3. 5. 3) the existence of many l-torsion classes in the cohomology of certain arithmetic subgroups of $SL(3, \mathbf{Z})$. The torsion classes we construct are Hecke eigenclasses whose eigenvalues are congruent to the Hecke eigenvalues of an automorphic (hence nontorsion) cohomology class in a different cohomology group. We wonder whether or not *all* torsion eigenclasses in the cohomology of any arithmetic group have this property. We also ask the related question: Is there a correspondence à la Langlands between l-torsion eigenclasses and Galois representations on $\mathbf{F}_l$-vector spaces?

1. 1 Cohomology and Hecke operators

In this section we define the Hecke algebra and state the basic properties of this algebra acting on the group cohomology. We will use the notation of Andrianov [1].

Let G be a group. A *Hecke pair* consists of a subgroup Γ of G and a subsemigroup S of G such that

(1) $\Gamma \subseteqq S$,

(2) Γ and $g^{-1}\Gamma g$ are commensurable for every $g \in S$.

We will write $L(\Gamma, S)$ for the free $\mathbf{Z}$-module on the right cosets Γg, $g \in S$, and $\mathscr{H} = \mathscr{H}(\Gamma, S)$ for the right Γ-invariant elements in $L(\Gamma, S)$. We define multiplication in $\mathscr{H}$ by the formula

$$\sum a_i(\Gamma g_i) \cdot \sum b_j(\Gamma h_j) = \sum a_i b_j(\Gamma g_i h_j).$$

Then $\mathscr{H}$ is an associative algebra. If $g \in S$ and $\Gamma g \Gamma$ is the disjoint union $\bigcup \Gamma g_i$ then we will write T_g for the element $\sum \Gamma g_i$ in $\mathscr{H}$. We will refer to $\mathscr{H}$ as the Hecke algebra of the pair (Γ, S).

Notation. If E is a right S-module, then we will denote the right action of $\sigma \in S$ on $e \in E$ by multiplication on the *left* by σ^{-1}:

$$E \times S \to E, \quad (e, \sigma) \mapsto \sigma^{-1} e.$$

It is well known that if E is a right $\mathbf{Z}S$-module, then there is a natural right action of the Hecke algebra $\mathscr{H}$ on the cohomology groups $H^r(\Gamma, E)$. For $g \in S$ the element T_g of $\mathscr{H}$ operates by the formula

$$(fT_g)(\gamma_0, \ldots, \gamma_r) = \sum g_i^{-1} f(t_i(\gamma_0), \ldots, t_i(\gamma_r)).$$

Here $f\colon \Gamma^{r+1} \to E$ is a homogeneous r-cocycle, $\gamma_0, \ldots, \gamma_r$ are in Γ, $\Gamma g \Gamma$ is the disjoint union $\bigcup \Gamma g_i$, and $t_i\colon \Gamma \to \Gamma$ is defined by the equations $\Gamma g_i \gamma = \Gamma g_j$ (for some j depending on i and $\gamma \in \Gamma$) and $g_i \gamma = t_i(\gamma) g_j$. The cohomology class of fT_g does not depend on the choice of the g_i.

Dually, if E is a left $\mathbf{Z}S$-module, then we can define a natural left action of $\mathscr{H}$ on the cohomology by formulas like the ones above. Alternatively, we set $S^{-1} = \{g^{-1} | g \in S\}$ and observe that since (Γ, S) is a Hecke pair, so also is (Γ, S^{-1}). As in the last paragraph we have a right action of $\mathscr{H}(\Gamma, S^{-1})$ on the cohomology. The left action of $\mathscr{H}$ is then given by the formula $T_g f = f T_{g^{-1}}$ for $g \in S$, and $f \in H^r(\Gamma, E)$. For more details see for instance [20].

It is fundamental to what follows that the action of the Hecke algebra on the cohomology groups respects the standard constructions of homological algebra. In the remainder of this section we record some instances of this principle.

Lemma 1.1.1. *Let $0 \to D \to E \to F \to 0$ be an exact sequence of (right or left) $\mathbf{Z}S$-modules. Then the long exact cohomology sequence*

$$\cdots \to H^r(\Gamma, D) \to H^r(\Gamma, E) \to H^r(\Gamma, F) \to H^{r+1}(\Gamma, D) \to \cdots$$

commutes with the action of $\mathscr{H}$.

Definition 1.1.2. A Hecke pair (Γ_0, S_0) is said to be compatible to the Hecke pair (Γ, S) if (a) $(\Gamma_0, S_0) \subseteq (\Gamma, S)$, (b) $\Gamma S_0 = S$, and (c) $\Gamma \cap S_0 S_0^{-1} = \Gamma_0$.

If $(\Gamma_0, S_0) \subseteq (\Gamma, S)$ are compatible, the cosets Γg with $g \in S_0$ span $L(\Gamma, S)$. So there is a unique linear map $L(\Gamma, S) \to L(\Gamma_0, S_0)$ sending Γg to $\Gamma_0 g$ for $g \in S_0$. The compatibility condition guarantees that the restriction to double cosets,

$$\iota\colon \mathscr{H} \to \mathscr{H}(\Gamma_0, S_0),$$

is an injective algebra homomorphism.

Thus if E is a $\mathbf{Z}S_0$-module, then we may (and will) view the cohomology groups $H^r(\Gamma_0, E)$ as $\mathscr{H}$-modules by composing the action of $\mathscr{H}(\Gamma_0, S_0)$ with ι.

The notion of compatibility (1. 1. 2) is not left-to-right symmetric. For this reason our next two results must treat right and left modules separately.

Lemma 1. 1. 3. (a) *If E is a right $\mathbf{Z}S$-module then the restriction map*

$$H^r(\Gamma; E) \xrightarrow{\text{res}} H^r(\Gamma_0, E)$$

commutes with the action of $\mathscr{H}$.

(b) *If the index $[\Gamma : \Gamma_0]$ is finite and E is a left $\mathbf{Z}S$-module then the corestriction map*

$$H^r(\Gamma, E) \xleftarrow{\text{cores}} H^r(\Gamma_0, E)$$

commutes with the action of $\mathscr{H}$.

Now suppose $[\Gamma : \Gamma_0] < \infty$ and let E be a right (resp. left) $\mathbf{Z}S_0$-module. Then in particular E is a $\mathbf{Z}\Gamma_0$-module and we may consider the induced $\mathbf{Z}\Gamma$-module $I = \mathrm{Ind}(\Gamma_0, \Gamma; E)$ of functions $f: \Gamma \to E$ such that $f(xy) = xf(y)$ for all x in Γ_0, y in Γ. We define a right (resp. left) action of S on I by the following formulas for $g \in S$, $f \in I$, and $x \in \Gamma$.

If E is a *right* S_0-module, choose $g_0 \in S_0$, $y \in \Gamma$ so that $xg^{-1} = g_0^{-1}y$ and set

$$(g^{-1}f)(x) = g_0^{-1}f(y).$$

If E is a *left* S_0-module, set

$$(gf)(x) = \sum xg\gamma^{-1}f(\gamma)$$

where the sum is over representatives, γ, of the cosets in $\Gamma_0\backslash(\Gamma \cap S_0^{-1}xg)$. Using the compatibility of $(\Gamma_0, S_0) \subseteqq (\Gamma, S)$ it is not difficult to verify that these formulas define a right (resp. left) semigroup action of S on I extending the standard action of Γ.

Lemma 1. 1. 4. *Suppose $[\Gamma : \Gamma_0] < \infty$ and let E be a (right or left) $\mathbf{Z}S_0$-module. Then the Shapiro isomorphism*

$$\mathscr{S}: H^r(\Gamma, \mathrm{Ind}(\Gamma_0, \Gamma, E)) \xrightarrow{\sim} H^r(\Gamma_0, E)$$

commutes with the action of $\mathscr{H}$.

Proof. Let $I = \mathrm{Ind}(\Gamma_0, \Gamma; E)$. If E is a right $\mathbf{Z}S$-module then the map $\rho: I \to E$ which sends a function f to $f(1)$ is a morphism of right S_0-modules. The Shapiro isomorphism is the composition

$$H^r(\Gamma, I) \xrightarrow{\text{res}} H^r(\Gamma_0, I) \xrightarrow{\rho_*} H^r(\Gamma_0, E),$$

and hence commutes with the action of $\mathscr{H}$ by 1. 1. 1 and 1. 1. 3.

If E is a left $\mathbf{Z}S_0$-module, then the map $i: E \to I$ defined by

$$i(e)(x) = \begin{cases} xe, & \text{if } x \in \Gamma_0, \\ 0, & \text{otherwise} \end{cases}$$

is a morphism of left $\mathbf{Z}S_0$-modules. The inverse of the Shapiro isomorphism is the composition

$$H^r(\Gamma_0, E) \xrightarrow{i_*} H^r(\Gamma_0, I) \xrightarrow{\text{cores}} H^r(\Gamma, I).$$

Again we use 1.1.1 and 1.1.3 to conclude that this commutes with $\mathscr{H}$. □

We close this section with a discussion of nebentype operators. Suppose (Γ_1, S_1) is a Hecke pair compatible to (Γ_0, S_0) and which is normalized by Γ_0. Let E be an S_0-module. We have seen that the Hecke algebra $\mathscr{H}_0 = \mathscr{H}(\Gamma_0, S_0)$ acts on $H^*(\Gamma_1, E)$. But there is also a standard action of the quotient group Γ_0/Γ_1 on the cohomology ([27], VII §§ 5, 6). In our language this action can be described as follows. Since Γ_1 is normal in Γ_0 the pair (Γ_1, Γ_0) is a Hecke pair. The Hecke algebra $\mathscr{H}(\Gamma_1, \Gamma_0)$ is naturally isomorphic to the group ring $\mathbf{Z}[\Gamma_0/\Gamma_1]$. Thus the action of $\mathscr{H}(\Gamma_1, \Gamma_0)$ on the cohomology induces an action of Γ_0/Γ_1 on $H^*(\Gamma_1, E)$. For $a \in \Gamma_0$ we write $[a]$ for the associated operator on cohomology and refer to $[a]$ as the nebentype operator associated to a. One readily verifies that these operators commute with the action of $\mathscr{H}_0$. Since the nebentype operators are defined as Hecke operators, they enjoy all of the functorial properties attributed to Hecke operators in this chapter.

Let R be a commutative ring with identity and let $\varepsilon: \Gamma_0/\Gamma_1 \to R^*$ be a character. For a right (resp. left) RS_0-module E we define $H^*(\Gamma_1, E)(\varepsilon)$ to be the submodule of ξ in $H^*(\Gamma_1, E)$ on which the nebentype operators act via ε: $\xi[a] = \varepsilon(a)^{-1}\xi$ (resp. $[a]\xi = \varepsilon(a)\xi$).

The compatibility of $(\Gamma_1, S_1) \subseteq (\Gamma_0, S_0)$ guarantees that there is a unique extension of ε to a character $\varepsilon: S_0 \to R^*$ which is trivial on S_1. Let R_ε be the rank one R-module on which S_0 acts via ε.

Lemma 1.1.5. *Let (Γ_1, S_1) be a Hecke pair compatible to (Γ_0, S_0) and normalized by Γ_0. Suppose the index $[\Gamma_0 : \Gamma_1]$ is finite and invertible in R. Then for every (left or right) RS_0-module E and every character $\varepsilon: \Gamma_0/\Gamma_1 \to R^*$, the restriction map induces an isomorphism of $\mathscr{H}_0$-modules*

$$H^*(\Gamma_0, E \otimes R_\varepsilon) \cong H^*(\Gamma_1, E)(\varepsilon^{-1}).$$

Proof. The map $R \to R_\varepsilon$, $r \mapsto r$ is an isomorphism of RS_1-modules and thus induces an isomorphism of $\mathscr{H}_0$-modules $H^*(\Gamma_1, E) \to H^*(\Gamma_1, E \otimes R_\varepsilon)$. A simple calculation with cocycles shows that the space $H^*(\Gamma_1, E)(\varepsilon^{-1})$ is mapped onto $H^*(\Gamma_1, E \otimes R_\varepsilon)^{\Gamma_0}$. The invertibility of $[\Gamma_0 : \Gamma_1]$ together with the Hochschild-Serre spectral sequence now show that this is isomorphic to $H^*(\Gamma_0, E \otimes R_\varepsilon)$ via the restriction map. If the action of S on E is a right action then lemma 1.1.3(a) shows that restriction commutes with $\mathscr{H}_0$ and we are done. Otherwise we use 1.1.3(b) and the fact that cores $\circ$ res $= [\Gamma_0 : \Gamma_1]$ which is invertible in R to complete the proof. □

1. 2 Systems of Hecke eigenvalues

In this section we discuss the functorial properties of systems of Hecke eigenvalues associated to eigenvectors in the cohomology groups $H^*(\Gamma, E)$. We will take a more abstract point of view and let $\mathscr{H}$ be an arbitrary *commutative* algebra.

Definition 1. 2. 1. (a) A system of eigenvalues of $\mathscr{H}$ with values in a commutative ring R is a set theoretic map $\Phi\colon \mathscr{H} \to R$.

(b) The system Φ is said to occur in the $R\mathscr{H}$-module A if there is a *nonzero* $a \in A$ such that $Ta = \Phi(T)a$ for all $T \in \mathscr{H}$. Such an a is called a Φ-eigenvector.

We will prove the following two propositions.

Proposition 1. 2. 2. *Suppose R is a discrete valuation ring (respectively field). Let A, B be $R\mathscr{H}$-modules, finitely generated over R and $f\colon A \to B$ be a surjective $R\mathscr{H}$-morphism. Let $\Phi\colon \mathscr{H} \to R$ be a system of eigenvalues occurring in B, and $v \in B$ be a Φ-eigenvector. Let $Q \subsetneqq R$ be a prime ideal in the support of Rv. Then there is a discrete valuation ring (respectively field) R' of finite type over R and a system $\Psi\colon \mathscr{H} \to R'$ occurring in $A \otimes_R R'$ such that $\Psi(T) \equiv \Phi(T) \pmod{Q'}$ for all $T \in \mathscr{H}$ where Q' is the unique prime ideal of R' for which $Q' \cap R = Q$.*

This proposition generalizes a lemma of Deligne and Serre ([9], Lemma 6. 11) which considers the special case where A is free over a discrete valuation ring R and B is the reduction of A modulo the maximal ideal.

If R is a local ring then we will use a bar to denote reduction modulo the maximal ideal. Thus, if $P \subsetneqq R$ is the maximal ideal then $\bar{R} = R/P$; if M is an R-module then $\bar{M} = M \otimes \bar{R}$; if $m \in M$ then $\bar{m} = m \otimes 1 \in \bar{M}$; and if $\Phi\colon \mathscr{H} \to R$ is a system of eigenvalues then $\bar{\Phi}\colon \mathscr{H} \to \bar{R}$ is the composition of Φ with the canonical projection $R \to \bar{R}$.

Proposition 1. 2. 3. *Suppose R is a discrete valuation ring. Let A be an $R\mathscr{H}$-module finitely generated over R. If $\Phi\colon \mathscr{H} \to R$ occurs in A then Φ occurs in $\bar{A}$.*

Remark. Proposition 1. 2. 2 states that a system of eigenvalues occurring in B may, after finite base extension, be "lifted" to a system occurring in A. Note, however, that an eigen*vector* in B need not lift to an eigenvector in A. Proposition 1. 2. 3 is also more subtle than it may seem at first. For example, a Φ-eigenvector $a \in A$ may reduce to zero modulo the maximal ideal. It is in general not even possible to solve the equation $rb = a$ for $r \in R$, $b \in A$ with $\bar{b}$ a Φ-eigenvector in $\bar{A}$.

In preparation for the proof of proposition 1. 2. 2 we state and prove two simple lemmas.

Lemma 1. 2. 4. *Let R be a discrete valuation ring (respectively field) and let A be an $R\mathscr{H}$-module, finitely generated over R. Then there is a discrete valuation ring (respectively field) R' finite over R such that $A' = A \otimes_R R'$ possesses an $R'\mathscr{H}$-stable filtration*

$$A' = A'_0 \supseteq A'_1 \supseteq \cdots \supseteq A'_s = 0$$

in which the successive quotients are cyclic R'-modules.

Proof. It suffices to show that A has an $\mathscr{H}$-eigenvector after some finite base extension. If R is a field this is well known.

If A has no nonzero R-torsion then we may reduce to the case where R is a field by tensoring with the quotient field of R. Otherwise there is a nonzero $R\mathscr{H}$-submodule $A_0 \subseteqq A$ which is annihilated by the maximal ideal P of R. The action of $\mathscr{H}$ on A_0 factors through an action of $\mathscr{H} \otimes \bar{R}$. Thus we are again reduced to the case where the base ring is a field. □

We will refer to an $R\mathscr{H}$-filtration $A = A_0 \supseteqq \cdots \supseteqq A_s = 0$ as a *cyclic* $R\mathscr{H}$-filtration if the successive quotients are cyclic R-modules. If the integer s is minimal among all such filtrations of A we will call the filtration a *minimal* cyclic $R\mathscr{H}$-filtration.

Lemma 1.2.5. *Let R be a discrete valuation ring or a field. Let*

$$A = A_0 \supseteqq A_1 \supseteqq \cdots \supseteqq A_s = 0$$

be a minimal cyclic $R\mathscr{H}$-filtration of A, and suppose $T_0 \in \mathscr{H}$ annihilates the quotient A_i/A_{i+1} for some $i = 0, \ldots, s-1$. Then there is a nonzero $a \in A$ such that $T_0 a = 0$.

Proof. We have $T_0 A_i \subseteqq A_{i+1}$. For $j = 0, \ldots, s-1$ let

$$A'_j = \begin{cases} T_0 A_j & \text{if } 0 \leqq j \leqq i, \\ T_0 A_i \cap A_{j+1} & \text{if } i < j. \end{cases}$$

Then $T_0 A = A'_0 \supseteqq A'_1 \supseteqq \cdots \supseteqq A'_{s-1} = 0$ is a cyclic $R\mathscr{H}$-filtration of $T_0 A$ of length $s-1$. Because of the minimality of s, $T_0 A$ and A are not isomorphic as $R\mathscr{H}$-modules. Thus T_0 does not act injectively on A. □

Proof of Proposition 1.2.2. By lemma 1.2.4 it suffices to prove the following:

(*) If A possesses a cyclic $R\mathscr{H}$-filtration then the proposition is true with

$$R' = R.$$

By replacing $\mathscr{H}$ by the R-algebra generated by the image of $\mathscr{H}$ in $\mathrm{End}_R(A)$ we may assume $\mathscr{H}$ is an R-subalgebra of $\mathrm{End}_R(A)$. Since A is a finitely generated R-module, so is $\mathscr{H}$. We will prove (*) by induction on the minimal number of algebra generators of $\mathscr{H}$ over the image of the structure morphism $R \to \mathscr{H}$. If $R \to \mathscr{H}$ is surjective there is nothing to prove. So we let $\mathscr{H}_1$ be a subalgebra of $\mathrm{End}_R(A)$ and make the following inductive assumption.

(1) If $\mathscr{H} = \mathscr{H}_1$ then (*) is valid.

Now suppose $\mathscr{H} = \mathscr{H}_1[T]$ for some $T \in \mathrm{End}_R(A)$. We will prove (*) in this situation by induction on the length of a minimal cyclic $R\mathscr{H}$-filtration of A. If A is a cyclic R-module then (*) is immediate. Thus we may make the following inductive assumption.

(2) $s > 1$ is an integer and (*) holds for every $\mathscr{H}$-module A which possesses a cyclic $R\mathscr{H}$-filtration of length less than s.

Let A be an $R\mathscr{H}$-module which possesses a minimal cyclic $R\mathscr{H}$-filtration

$$A = A_0 \supseteq A_1 \supseteq \cdots \supseteq A_s = 0$$

of length s. Without loss of generality we may also make the following assumption on B.

(3) B is the cyclic R-module generated by v.

Let $\lambda = \Phi(T)$ and let C be the submodule of A defined by

$$C = \bigcup \ker \prod_{j=1}^{n} (T - \lambda_j)$$

where the union is over all natural numbers n and n-tuples $(\lambda_1, \ldots, \lambda_n) \in R^n$ whose components satisfy the congruence $\lambda_j \equiv \lambda \pmod{Q}$. Then C is $\mathscr{H}$-stable since $\mathscr{H}$ is commutative.

Suppose $f(C) = 0$. Then $f\colon A \to B$ factors through an $R\mathscr{H}$-morphism

$$f'\colon A' = A/C \to B.$$

Let $A' = A'_0 \supseteq A'_1 \supseteq \cdots \supseteq A'_s = 0$ be the $R\mathscr{H}$-filtration of A' induced by the given filtration of A. Let $0 \leqq i \leqq s-1$ be the least integer for which $f'(A'_{i+1}) = 0$. Let $\lambda' \in R$ satisfy $Tx = \lambda' x$ for all $x \in A'_i/A'_{i+1}$. Since f' induces a nonzero $R\mathscr{H}$-morphism $A'_i/A'_{i+1} \to B$ and B is cyclic (3), we have $\lambda' \equiv \lambda \pmod{Q}$. By the definition of C we know that $T - \lambda'$ acts injectively on A'. On the other hand $T - \lambda'$ annihilates A'_i/A'_{i+1}. Thus lemma 1.2.5 implies the filtration $A' = A'_0 \supseteq \cdots \supseteq 0$ is not minimal cyclic. Then by (2) there is a system of eigenvalues $\Psi'\colon \mathscr{H} \to R$ occurring in A' such that $\Psi' \equiv \Phi \pmod{Q}$. But then $\Psi'(T) \equiv \lambda \pmod{Q}$ and therefore, by the definition of C, $T - \Psi'(T)$ acts injectively on A', a contradiction.

This shows $f(C) \neq 0$. By applying (1) to the surjection $C \to f(C)$ we conclude that there is a system $\Psi_1\colon \mathscr{H}_1 \to R$ occurring in C such that $\Psi_1(t) \equiv \Phi(t) \pmod{Q}$ for all $t \in H_1$. Let $c \in C$ be a Ψ_1-eigenvector. By the definition of C there is a nonnegative integer n and $\lambda_1, \ldots, \lambda_n \in R$ with $\lambda_j \equiv \lambda \pmod{Q}$, $j = 1, \ldots, n$ such that $a = \prod_{j=2}^{n} (T - \lambda_j) c \neq 0$, and $(T - \lambda_1) a = 0$. Then a is a Ψ-eigenvector for a system of eigenvalues $\Psi\colon \mathscr{H} \to R$ satisfying $\Psi \equiv \Phi \pmod{Q}$. □

Proof of Proposition 1.2.3. As in the proof of proposition 1.2.2 we may suppose $\mathscr{H} \subseteq \mathrm{End}_R(A)$ and $\mathscr{H} = \mathscr{H}_1[T]$ where $\Phi|\mathscr{H}_1$ occurs in $\bar{A}$. Without loss of generality we may replace A by the full inverse image of the $\Phi|\mathscr{H}_1$-eigenspace of $\bar{A}$, since Φ occurs already in this submodule. Thus it suffices to find a nonzero $x \in \bar{A}$ such that $Tx = rx$ where $r = \Phi(T)$.

Since A is a finite direct sum of cyclic R-modules, there is a free R-module F and a surjective R-morphism $\xi\colon F \to A$ whose reduction $\bar{\xi}\colon \bar{F} \to \bar{A}$ is an isomorphism. We may lift T to an R-endomorphism of F so that T commutes with ξ.

By proposition 1.2.2 there is a discrete valuation ring R' finite over R, and a nonzero $f \in F \otimes_R R'$ such that $Tf = sf$ for some $s \in R'$ with $\bar{s} = \bar{r}$. Since F is free we may *assume* $\bar{f} \neq 0$. Thus $(t - \bar{r})$ has a nonzero kernel in $\bar{F} \otimes_R \bar{R}' \cong \bar{A} \otimes_R \bar{R}'$, and hence also in $\bar{A}$. □

1.3 The main diagrams

We consider compatible Hecke pairs $(\Gamma_0, S_0) \subseteq (\Gamma, S)$ and let $\mathscr{H}$ be the associated Hecke algebra. We assume the following additional conditions:

(1.3.1) Γ is finitely presented and of type (WFL) ([26], section 1.8), $[\Gamma : \Gamma_0] < \infty$.

Let R be a discrete valuation ring with maximal ideal P generated by π, and let E be a right (left) RS-module and F be a right (left) RS_0-module. We assume that both E and F are finitely generated as R-modules. This assumption taken together with (1.3.1) implies that the cohomology groups of E and F are finitely generated over R as well, so that we can apply the results of the last section to them.

If E and F are right modules and $\phi: \bar{E} \to \bar{F}$ is an RS_0-morphism then the map $\alpha(\phi): \bar{E} \to \mathrm{Ind}(\Gamma_0, \Gamma; \bar{F})$ defined by $(\alpha(\phi)(e))(\gamma) = \phi(\gamma e)$ for $e \in \bar{E}$ and $\gamma \in \Gamma$ commutes with the right action of RS. The interesting cases of course occur when ϕ cannot be lifted to a morphism $E \to F$.

For a positive integer r we can now draw the main diagram for *right* modules.

(1.3.2)
$$\begin{array}{ccccc}
H^r(\Gamma, E) & \longrightarrow & H^r(\Gamma, \bar{E}) & \xrightarrow{\delta} & H^{r+1}(\Gamma, E) \\
 & \overset{\mathrm{res}}{\searrow} & \downarrow A^r & \overset{\alpha(\phi)_*}{\swarrow} & \\
H^r(\Gamma_0, \bar{E}) & & & & H^r(\Gamma, \mathrm{Ind}(\Gamma_0, \Gamma; \bar{F})) \\
 & \underset{\phi_*}{\searrow} & & \underset{\mathscr{S}}{\swarrow} & \\
H^r(\Gamma_0, F) & \longrightarrow & H^r(\Gamma_0, \bar{F}) & \xrightarrow{\delta} & H^{r+1}(\Gamma_0, F)
\end{array}$$

In this diagram the horizontal arrows are extracted from the long exact cohomology sequences of $0 \to M \xrightarrow{\pi} M \to \bar{M} \to 0$ for $M = E, F$, the arrow res is the restriction morphism, and $\mathscr{S}$ is the Shapiro isomorphism. The arrow A^r is defined by the commutativity of the diagram. The results of 1.1 show that this diagram commutes with the action of $\mathscr{H}$.

Dually, if E and F are left modules and $\psi: \bar{F} \to \bar{E}$ is an RS_0-morphism then we define $\beta(\psi): \mathrm{Ind}(\Gamma_0, \Gamma; \bar{F}) \to \bar{E}$ by

$$\beta(\psi)(f) = \sum_{\gamma \in \Gamma_0 \backslash \Gamma} \gamma^{-1} \psi(f(\gamma)).$$

A straightforward calculation shows that $\beta(\psi)$ is a morphism of left RS-modules.

Our main diagram for *left* modules is the following.

(1.3.3)
$$\begin{array}{ccccc}
H^r(\Gamma, E) & \longrightarrow & H^r(\Gamma, \bar{E}) & \xrightarrow{\delta} & H^{r+1}(\Gamma, E) \\
 & \overset{\mathrm{cores}}{\nearrow} & \downarrow B^r & \overset{\beta(\psi)_*}{\nwarrow} & \\
H^r(\Gamma_0, \bar{E}) & & & & H^r(\Gamma, \mathrm{Ind}(\Gamma_0, \Gamma; \bar{F})) \\
 & \underset{\psi_*}{\nearrow} & & \underset{\mathscr{S}^{-1}}{\nwarrow} & \\
H^r(\Gamma_0, F) & \longrightarrow & H^r(\Gamma_0, \bar{F}) & \xrightarrow{\delta} & H^{r+1}(\Gamma_0, F)
\end{array}$$

Here cores is the corestriction morphism which commutes with the action of $\mathscr{H}$ by lemma 1. 1. 3, and B^r is defined by the commutativity of the diagram.

Theorem 1. 3. 4. *Suppose $\mathscr{H}$ is commutative and replace R by a finite extension if necessary so that R contains all of the eigenvalues of $\mathscr{H}$ acting on the cohomology groups in* (1. 3. 2) *and* (1. 3. 3). *Let $\Phi\colon \mathscr{H} \to R$ be a system of eigenvalues.*

(a) *Suppose A^r is injective or B^r is surjective. If Φ occurs in $H^r(\Gamma, E)$ then there is a system of eigenvalues $\Psi\colon \mathscr{H} \to R$ occurring in $H^r \oplus H^{r+1}(\Gamma_0, F)$ such that $\bar{\Phi} = \bar{\Psi}$.*

(b) *Suppose A^r is surjective or B^r is injective. If Φ occurs in $H^r(\Gamma_0, F)$ then there is a system of eigenvalues $\Psi\colon \mathscr{H} \to R$ occurring in $H^r \oplus H^{r+1}(\Gamma, E)$ such that $\bar{\Phi} = \bar{\Psi}$.*

Proof. We will give the proof of (a) only. The proof of (b) is similar. Suppose Φ occurs in $H^r(\Gamma, E)$. The long exact cohomology sequence of the sequence $0 \to E \xrightarrow{\pi} E \to \bar{E} \to 0$ provides a Hecke equivariant inclusion

$$H^r(\Gamma, E) \otimes R/P \hookrightarrow H^r(\Gamma, \bar{E}).$$

Thus proposition 1. 2. 3 shows that $\bar{\Phi}$ occurs in $H^r(\Gamma, \bar{E})$. If A^r is injective we conclude at once that $\bar{\Phi}$ occurs in $H^r(\Gamma_0, \bar{F})$, and if B^r is surjective we use proposition 1. 2. 2 to draw the same conclusion. In either case we can find a $\bar{\Phi}$-eigenvector $v \in H^r(\Gamma_0, F)$. If $\delta(v) \neq 0$ then Φ occurs in $H^{r+1}(\Gamma_0, F)$. Otherwise we can appeal to proposition 1. 2. 2 again to prove the existence of Ψ occurring in $H^r(\Gamma_0, F)$ with $\bar{\Phi} = \bar{\Psi}$. □

We close this section by giving a criterion for A^r or B^r to be surjective.

Theorem 1. 3. 5. *Let N be the virtual cohomological dimension of Γ and D be the greatest common divisor of the indices of the torsionfree subgroups of finite index in Γ. Assume D is invertible in R. Then we have the following implications:*

(i) $\alpha(\phi)$ *surjective* $\Rightarrow$ A^N *surjective*;

(ii) $\beta(\psi)$ *surjective* $\Rightarrow$ B^N *surjective.*

Proof. To prove (i) we consider the long exact cohomology sequence associated to

$$0 \to \ker(\alpha) \to \bar{E} \to \mathrm{Ind}(\Gamma_0, \Gamma; \bar{F}) \to 0.$$

By [6], p. 287 we have $D \cdot H^{N+1}(\Gamma, \ker(\alpha)) = 0$. Because D is invertible in R we have $H^{N+1} = 0$ and therefore $\alpha_*\colon H^N(\Gamma, \bar{E}) \to H^N(\Gamma, \mathrm{Ind})$ is surjective. Since $A^N = \mathscr{S} \circ \alpha_*$ and $\mathscr{S}$ is an isomorphism, (i) follows. A similar argument establishes (ii). □

1.4 Arithmetic groups

In the applications Γ will be an arithmetic subgroup of a reductive algebraic group G over $\mathbf{Q}$, and S will be a subsemigroup of $G(\mathbf{Q})$. Then Γ satisfies (1.3.1) and the results of the last section are applicable.

The additional structure which we would like to utilize derives from the fact that Γ acts properly discontinuously on the symmetric space X of G. We define the integers d, N, m and D by

$$\begin{aligned} d &= \text{the dimension of } X; \\ N &= \text{the virtual cohomological dimension of } \Gamma; \\ m &= \text{the least common multiple of the orders } |\Gamma_x| \\ &\quad \text{of the isotropy groups of all } x \text{ in } X; \\ D &= \text{the greatest common divisor of the indices of} \\ &\quad \text{the torsionfree subgroups of finite index in } \Gamma. \end{aligned} \tag{1.4.1}$$

Borel and Serre [4] have shown that N depends only on G and is equal to d minus the $\mathbf{Q}$-rank of G.

Let R be a ring in which m is invertible. For an $R\Gamma$-module E, let $\tilde{E}_\Gamma$ be the corresponding local coefficient system on the quotient $X_\Gamma = \Gamma\backslash X$, of X by Γ. The invertibility of m implies that the canonical map

$$H^r(X_\Gamma, \tilde{E}_\Gamma) \longrightarrow H^r(\Gamma, E)$$

is an isomorphism for every r. Letting H_c denote cohomology with compact supports and $H_!$ the image of H_c in H, we define

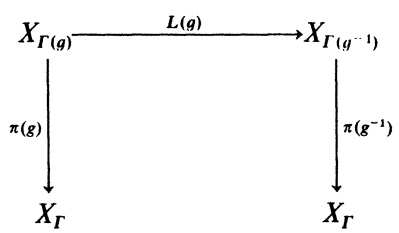
$$H_c^r(\Gamma, E) = H_c^r(X_\Gamma, \tilde{E}_\Gamma) \quad \text{and} \quad H_!^r(\Gamma, E) = H_!^r(X_\Gamma, \tilde{E}_\Gamma).$$

If E is a right (resp. left) RS-module, then the action of $\mathscr{H}$ can be described topologically. For g in S let

$$\Gamma(g) = \Gamma \cap g^{-1}\Gamma g \quad \text{and} \quad \Gamma(g^{-1}) = g\Gamma(g)g^{-1} = \Gamma \cap g\Gamma g^{-1},$$

and consider the diagram

(1.4.2)



where $\pi(g)$ and $\pi(g^{-1})$ are the natural projections and $L(g)$ is induced by left translation of X by g. Then $T_g\colon H_*^r(\Gamma, E) \to H_*^r(\Gamma, E)$ is given by

$$T_g = \pi(g)_* \circ L(g)^* \circ \pi(g^{-1})^*$$

(resp. $T_g = \pi(g^{-1})_* \circ L(g)_* \circ \pi(g)^*$) where H_* denotes either H or H_c. The map $H_c^r \to H^r$ commutes with the action of $\mathscr{H}$ so that $H_!^r(\Gamma, E)$ also inherits a structure of $\mathscr{H}$-module.

If E is a right (resp. left) RS-module finitely generated as an R-module, we define the contragredient left (resp. right) RS-module $E^* = \mathrm{Hom}_R(E, R)$ by $(gf)(e) = f(g^{-1}e)$ (resp. $(g^{-1}f)(e) = f(ge)$) for $f \in E^*$, $e \in E$, and $g \in S$.

The following lemma will not be used in the rest of the paper, though it seems appropriate to record it here for the sake of completeness.

Lemma 1.4.3. *Suppose Γ acts on X without reversing orientation, and let R be a field in which mD is invertible. Then cup product and the identification $H_c^d(\Gamma, R) = R$ induce $\mathscr{H}$-equivariant perfect pairings*

(i) $H_c^r(\Gamma, E) \times H^{d-r}(\Gamma, E^*) \to R$

and

(ii) $H_!^r(\Gamma, E) \times H_!^{d-r}(\Gamma, E^*) \to R$.

Proof. If Γ is torsion free then the Poincaré duality theorem ([7], p. 20—40) assures that the pairing (i) is nondegenerate. In the general case let Γ' be a torsion free subgroup of finite index in Γ such that $[\Gamma : \Gamma']$ is invertible in R. Let $\pi: X_{\Gamma'} \to X_\Gamma$ be the canonical projection. For each $R\Gamma$-module M and each integer $r \geqq 0$ we consider the maps

$$\pi^*: H_*^r(\Gamma, M) \to H_*^r(\Gamma', M), \quad \pi_*: H_*^r(\Gamma, M) \leftarrow H_*^r(\Gamma', M)$$

where H_* denotes either H or H_c. Since $\pi_* \circ \pi^*$ is multiplication by $[\Gamma : \Gamma']$ which is invertible in R we see that π_* is surjective and π^* is injective. If $x \in H_c^r(\Gamma, E)$ satisfies $\langle x, y \rangle = 0$ for every $y \in H^{d-r}(\Gamma, E^*)$ then $0 = \langle x, \pi_*(z) \rangle = \langle \pi^*(x), z \rangle$ for all $z \in H^{d-r}(\Gamma', E^*)$. Since Γ' is torsion free we see that $\pi^*(x) = 0$. But π^* is injective, so $x = 0$. Thus the pairing is nondegenerate on the left. The right nondegeneracy is proved similarly.

Next we check that (i) is $\mathscr{H}$-equivariant. To fix ideas we suppose E is a *right* RS-module. Let $x \in H_c^r(\Gamma, E)$ and $y \in H^{d-r}(\Gamma, E^*)$. For $g \in S$ we have

$$\langle xT_g, y \rangle = \langle \pi(g)_* \circ L(g)^* \circ \pi(g^{-1})^* x, y \rangle = \langle x, \pi(g^{-1})_* \circ L(g)_* \circ \pi(g)^* y \rangle = \langle x, T_g y \rangle.$$

The canonical maps $H_c^*(\Gamma, E) \to H^*(\Gamma, E)$ and $H_c^*(\Gamma, E^*) \to H^*(\Gamma, E^*)$ are dual under the pairing (i) and commute with $\mathscr{H}$. Thus (ii) is a consequence of (i). □

2. Systems of Hecke eigenvalues mod l

In this section, we prove two general theorems (2.2 and 2.4) about systems of Hecke eigenvalues (mod l) occurring in the cohomology groups of a fixed arithmetic group Γ. As mentioned in the introduction these results generalize known statements about classical modular forms [19], [29], [31].

We begin by defining precisely the objects to concern us for the rest of this paper. Let l be a rational prime and let $\mathbf{Z}_{(l)}$ be the ring of rational numbers with denominators prime to l. We make the following assignments.

G = a reductive linear algebraic group scheme defined over $\mathbf{Z}_{(l)}$.

Γ = an arithmetic group in $G(\mathbf{Z}_{(l)})$.

(Γ, S) = a Hecke pair.

$\mathscr{H}$ = the Hecke algebra $\mathscr{H}(\Gamma, S)$.

$\mathscr{E}$ = a finite dimensional representation of G defined over $\mathbf{Z}_{(l)}$.

E = a Γ-stable finitely generated $\mathbf{Z}$-module in $\mathscr{E}(\mathbf{Z}_{(l)})$ such that $\mathbf{Z}_{(l)}E = \mathscr{E}(\mathbf{Z}_{(l)})$.

We will call such an E an l-rational $\mathbf{Z}\Gamma$-module. We also let

$\mathcal{O}$ = the ring of algebraic integers in the algebraic closure $\bar{\mathbf{Q}}$ of $\mathbf{Q}$.

λ = a prime ideal in $\mathcal{O}$ lying over l.

We will identify $\mathcal{O}/\lambda$ with the algebraic closure of the finite field with l elements via a *fixed* isomorphism $\mathcal{O}/\lambda \cong \bar{F}_l$. If we reduce an integral object mod l or λ, we use a bar to denote the result. The completion of $\mathcal{O}$ at λ will be denoted by $\mathcal{O}_\lambda$.

Lemma 2. 1. *Suppose $\mathscr{H}$ is commutative. Let k be a field, and V be a (left or right) kS-module, finite dimensional over k. If $\Phi\colon \mathscr{H} \to k$ occurs as a system of eigenvalues in $H^r(\Gamma, V)$, then Φ occurs in $H^r(\Gamma, W)$ for some irreducible kS-subquotient W of V.*

Proof. If V is not already irreducible, we have an exact sequence of kS-modules $0 \to V_1 \to V \to V_2 \to 0$ with nonzero V_1 and V_2. We get an exact sequence of $\mathscr{H}$-modules $H^r(\Gamma, V_1) \to H^r(\Gamma, V) \to H^r(\Gamma, V_2)$. If a Φ-eigenvector maps nonzero into $H^r(\Gamma, V_2)$, Φ occurs in the latter. Otherwise, this eigenvector is in the image of $H^r(\Gamma, V_1)$ and Φ occurs in $H^r(\Gamma, V_1)$ by proposition 1. 2. 2. The result follows from induction on $\dim_k V$. □

Theorem 2. 2. *Suppose the Hecke algebra $\mathscr{H}$ is commutative. Given an integer r, form the set $Z = \{\bar{\Phi}\colon \mathscr{H} \to \bar{F}_l\}$ with Φ running through all systems of eigenvalues $\Phi\colon \mathscr{H} \to \mathcal{O}_\lambda$ occurring in $\oplus H^r(\Gamma, E \otimes \mathcal{O}_\lambda)$, where E runs over all l-rational $\mathbf{Z}\Gamma$-modules. Then Z is finite.*

Proof. Let $J \subseteq G(\mathbf{F}_l)$ be the reduction of S mod l. Then J is a finite semigroup with cancellation laws and is therefore a finite group. Let $\{V\}$ be a set of representatives of isomorphism classes of irreducible finitely generated $\bar{F}_l J$-modules. This is a finite set (see, e.g. Part III of [30]). Define the set $Z_1 = \{\Psi\colon \mathscr{H} \to \bar{F}_l\}$ where Ψ runs through all systems of eigenvalues occurring in $\oplus H^r(\Gamma, V)$, V running over $\{V\}$. Each $H^r(\Gamma, V)$ is finite-dimensional over $\bar{F}_l$, so Z_1 is a finite set. We will show that Z is contained in Z_1.

Let E and Φ be as in the statement of the theorem. By lemma 1. 2. 4 there is a discrete valuation ring $R \subseteq \bar{\mathbf{Q}}$ of finite type over $\mathbf{Z}_{(l)}$ and with maximal ideal $R \cap \lambda$, such that Φ takes values in R and occurs in $H^r(\Gamma, E \otimes R)$. Then from proposition 1. 2. 3 it follows that $\bar{\Phi}$ occurs in $H^r(\Gamma, E \otimes R) \otimes R/(\lambda \cap R)$. From the Bockstein exact sequence we know this last injects into $H^r(\Gamma, E \otimes R/(\lambda \cap R))$, which injects into $H^r(\Gamma, E \otimes \bar{\mathbf{F}}_l)$. So $\bar{\Phi}$ occurs in the latter.

Now S acts on $E \otimes \bar{\mathbf{F}}_l$ via the reduction map $S \to J$. By lemma 2.1, we see that Φ occurs in $H^r(\Gamma, V)$ for some V in $\{V\}$. So $\bar{\Phi}$ is in Z_1. □

Let E be an l-rational $\mathbf{Z}\Gamma$-module and e be a nonzero element of $\bar{E}$. Set

$$S(e) = \{g \in G(\mathbf{Z}_{(l)}) \mid ge \in \Gamma e\},$$
$$S_1(e) = \{g \in S(e) \mid ge = e\},$$
$$\Gamma_0 = \{g \in \Gamma \mid ge \in \mathbf{F}_l^* e\}.$$

Proposition 2.3. *Let S be a subsemigroup of $S(e)$ which contains Γ and set $\Gamma_1 = \Gamma \cap S_1(e)$, $S_1 = S \cap S_1(e)$ and $S_0 = \Gamma_0 S_1$. Then $(\Gamma_1, S_1) \subseteqq (\Gamma_0, S_0) \subseteqq (\Gamma, S)$ are compatible Hecke pairs and Γ_0 normalizes (Γ_1, S_1).*

Proof. Clearly Γ_1 and Γ_0 have finite index in Γ, so Γ_1, Γ_0 are arithmetic, and $G(\mathbf{Z}_{(l)})$ commensurates them. Thus (Γ_1, S_1), (Γ_0, S_0) and (Γ, S) are Hecke pairs.

We will prove only the compatibility of (Γ_1, S_1) to (Γ, S). The other two compatibility statements are easy consequences of this one. We must verify conditions (b) and (c) of definition 1.1.2. Clearly $\Gamma S_1 \subseteqq S$. If $g \in S$ then $ge = \gamma e$ for some $\gamma \in \Gamma$. Thus $\gamma^{-1} g \in S_1$ and it follows that $g \in \Gamma S_1$ proving (b). If $\gamma = gh^{-1} \in \Gamma$ with $g, h \in S_1$ then

$$\gamma e = (gh^{-1})e = (gh^{-1})he = ge = e,$$

so $\gamma \in \Gamma_1$. This shows $\Gamma \cap S_1 S_1^{-1} = \Gamma_1$ which is (c). □

Theorem 2.4. *Let N, D be as in (1.4.1) and suppose l does not divide D. Let S, Γ_1, Γ_0 be as in proposition 2.3 and $\varepsilon: \Gamma_0/\Gamma_1 \to \mathbf{Z}_l^*$ be the unique character satisfying $ae = \varepsilon(a)e$ for all a in Γ_0.*

If $\mathscr{H}$ is commutative and $\bar{E}$ is irreducible as an $\mathbf{F}_l\Gamma$-module, then for each system of eigenvalues $\Phi: \mathscr{H} \to \mathcal{O}_\lambda$ occurring in $H^N(\Gamma, E \otimes \mathcal{O}_\lambda)$ there is a system $\Phi_1: \mathscr{H} \to \mathcal{O}_\lambda$ occurring in $H^N(\Gamma_1, \mathcal{O}_\lambda)(\varepsilon^{-1})$ such that $\bar{\Phi} = \bar{\Phi}_1$.

Proof. Let S_1, S_0 be as in proposition 2.3 and let $\varepsilon: S_0 \to \mathbf{Z}_l^*$ be the unique extension of ε which is trivial on S_1. Then the map $\psi: (\mathbf{F}_l) \to \bar{E}$ defined by $\psi(r) = re$ is S_0-equivariant by the definition of ε. We can therefore draw diagram 1.3.3 with $F = (\mathbf{Z}_l)_\varepsilon$ and $\bar{F} = (\mathbf{F}_l)_\varepsilon$.

The image of $\beta(\psi)$: $\mathrm{Ind}(\Gamma_0, \Gamma; (\mathbf{F}_l)_\varepsilon) \to \bar{E}$ contains e and is therefore a nonzero $\mathbf{F}_l\Gamma$-submodule of $\bar{E}$. Since $\bar{E}$ is irreducible we see that $\beta(\psi)$ is surjective. Theorem 1.3.5 (ii) tells us that B is surjective. Since $H^{N+1}(\Gamma_0, (\mathbf{Z}_l)_\varepsilon)$ vanishes, theorem 1.3.4 (a) shows that there is a system of eigenvalues Φ_1 occurring in $H^N(\Gamma_0, (\mathbf{Z}_l)_\varepsilon)$ for which $\bar{\Phi}_1 = \bar{\Phi}$. The theorem is now a consequence of lemma 1.1.5. □

Remark. For the sake of simplicity this result has been stated only for the case when $\bar{E}$ is irreducible. However a similar result is true for reducible $\bar{E}$ if we replace e by an element of some irreducible subquotient of $\bar{E}$ and define S, Γ_1, Γ_0 as before.

We close this section with an investigation of the irreducible representations of $G = GL(n)$. We first give a few preliminary definitions and remarks.

Let T be a Young tableau with $r=r(T)$ rows of lengths $g_1 \geqq \cdots \geqq g_r > 0$. Let $g=g(T)=g_1+\cdots+g_r$ and number the positions of T from 1 to g in lexicographic order. Let $\mathscr{S}_g$ be the symmetric group on $\{1,\ldots,g\}$. As in [36], Chapter IV, there is a unique idempotent $C_T \in \mathbf{Q}\mathscr{S}_g$ such that

$$pC_T = C_T, \quad C_T q = \operatorname{sgn}(q) C_T$$

for all $p, q \in \mathscr{S}_g$ such that p preserves the rows of T and q preserves the columns of T. In fact

$$C_T = \frac{1}{\mu_T} \sum_{p,q} \operatorname{sgn}(q) pq$$

where the sum is over all p, q as above and μ_T is a positive integer which divides $g!$.

Now let $G=GL(n)$ and $W=\mathbf{Q}^n$ be the standard representation of G over $\mathbf{Q}$. The symmetric group $\mathscr{S}_g$ acts on $\otimes^g W$ by permuting the factors. Let $W_T = C_T \cdot \otimes^g W$. As in [36], theorem 4.4F, it follows that a finite dimensional rational representation V of G is irreducible if and only if there is an integer v and a Young tableau T with $r(T)<n$ such that $V \cong W_T \otimes \det(\ )^v$.

We fix a tableau T with $r=r(T)<n$. We assume that the prime l is greater than $g=g(T)$, and as usual denote reduction modulo l by bar. Let $M=\mathbf{Z}^n \subseteqq W$ be the standard lattice and set $E_T = C_T \cdot \otimes^g M \subseteqq W_T$. Then M and E_T are l-rational $\mathbf{Z}\Gamma$-modules for every arithmetic group $\Gamma \subseteqq SL(n, \mathbf{Z})$.

Next we fix Γ to be $SL(n, \mathbf{Z})$. Then the reduction map $\Gamma \to SL(n, \mathbf{F}_l)$ is surjective and our assumption $l>g$ implies $\bar{E}_T$ is irreducible as an $\mathbf{F}_l\Gamma$-module. Define $\Gamma_1(l)$ to be the group of all $\gamma \in \Gamma$ which are congruent modulo l to an upper triangular matrix with ones on the diagonal.

We will identify M with the space of n-dimensional column vectors over $\mathbf{F}_l$ with the usual left action of $G(\mathbf{Z}_{(l)})$. Let $e_1 = {}^t(1, 0, \ldots, 0), \ldots, e_n = {}^t(0, \ldots, 0, 1)$ be the standard basis of $\bar{M}$ and set

$$e_T = C_T \cdot \left((\otimes^{g_1} e_1) \otimes \cdots \otimes (\otimes^{g_r} e_r)\right) \in \bar{E}_T.$$

One can check that $\Gamma_1(l)$ fixes e_T.

We take for S the set of all $\sigma \in M_n(\mathbf{Z})$ with $l \nmid \det(\sigma)$. Then $S \subseteqq S(e)$. Transposition defines an anti-isomorphism of S which leaves invariant the double cosets $\Gamma\sigma\Gamma$. Thus $\mathscr{H} = \mathscr{H}(\Gamma, S)$ is *commutative* ([32], proposition 3.8).

Proposition 2.5. *Let V be an irreducible finite dimensional rational representation of $GL(n, \mathbf{Q})$, and let $\Phi: \mathscr{H} \to \mathcal{O}$ be a system of eigenvalues occurring in $H^N(\Gamma, V)$ with $N = \frac{n(n-1)}{2}$. Then for sufficiently large l and every $\lambda \subseteqq \mathcal{O}$ extending l, there is a group Γ_1 intermediate to $\Gamma_1(l) \subseteqq \Gamma$ and a system of eigenvalues Φ_1 occurring in $H^N(\Gamma_1, \mathcal{O}_\lambda)$ such that $\bar{\Phi} = \bar{\Phi}_1$.*

Proof. The integer N in the proposition is the virtual cohomological dimension of Γ. Let m and D be as in (1.4.1). Let T be a Young tableau and $v \geqq 0$ be such that $V \cong W_T \otimes \det(\ )^v$. The conclusion of the proposition then follows from theorem 2.4 if we let l be greater than $g(T)$ and relatively prime to mD, and take Γ_1 to be the stabilizer of e_T. □

3. An example

In this section we combine the results of the first section with the theory of automorphic representations to prove the existence of torsion classes in the cohomology of certain subgroups of $SL(3, \mathbf{Z})$ (see theorem 3. 5. 3). To prove this theorem we need three lemmas which in principle are "well known". The first (proposition 3. 2. 1) states that any system of Hecke eigenvalues occurring in the cohomology of the boundary of the Borel-Serre compactification corresponds to a *reducible* Galois representation. The second (lemma 3. 3. 2) is a vanishing statement for interior cohomology with coefficients in a non self-dual representation. The third (lemma 3. 4. 3) concerns the existence in cohomology of the symmetric squares lift of a classical modular eigenform.

All of this is most easily understood in an adelic setting. Thus we begin by adelizing the cohomology groups and their associated Hecke algebras.

Notation. We will write $\mathbf{A}$ for the adeles of $\mathbf{Q}$ and $\mathbf{A}_f$ for the finite adeles. If B is an adelic object then we let B_∞, B_f, B_p be the infinite, respectively finite, respectively p-adic components of B. For example, if $g \in GL(3, \mathbf{A})$ then $g_\infty \in GL(3, \mathbf{R})$, $g_f \in GL(3, \mathbf{A}_f)$, and $g_p \in GL(3, \mathbf{Q}_p)$. Similarly, if $\chi: \mathbf{A}^* \to \mathbf{C}^*$ then $\chi_\infty: \mathbf{R}^* \to \mathbf{C}^*$, $\chi_f: \mathbf{A}_f^* \to \mathbf{C}^*$, and $\chi_p: \mathbf{Q}_p^* \to \mathbf{C}^*$.

3. 1 Adelization of the cohomology

Let $G_3 = GL(3)$, Z be the center of G_3, $Z_\infty = Z(\mathbf{R})$, and $K_\infty = SO(3)$. Then $X = G_3(\mathbf{R})/Z_\infty K_\infty$ is the symmetric space of G_3. To each compact open subgroup $K_f \subseteqq G_3(\mathbf{A}_f)$ we associate the topological space

$$X_{K_f} = G_3(\mathbf{Q}) \backslash G_3(\mathbf{A}) / Z_\infty K_\infty K_f .$$

This space has finitely many connected components and moreover, there are arithmetic subgroups $\Gamma_i \subseteqq SL(3, \mathbf{R})$, $i = 1, \ldots, h$, such that X_{K_f} is the disconnected union

$$X_{K_f} \cong \bigcup_{i=1}^{h} \Gamma_i \backslash X.$$

Let E be a finite dimensional irreducible representation over $\mathbf{C}$ of the group $G_3(\mathbf{R})$. The action will be a *left* action. We describe a sheaf $\tilde{E}_{K_f}$ over X_{K_f} by describing its local sections. Let $\pi: G_3(\mathbf{A})/Z_\infty K_\infty K_f \to X_{K_f}$ be the natural projection and let U be an open subset of X_{K_f}. Then

$$\tilde{E}_{K_f}(U) = \left\{ s: \pi^{-1}(U) \to E \;\middle|\; \begin{array}{l} s \text{ is locally constant and} \\ \text{for all } g \in G_3(\mathbf{Q}) \text{ and } x \in \pi^{-1}(U) \\ s(gx) = g_\infty \cdot s(x) \end{array} \right\}.$$

We can now form the cohomology groups $H^*(X_{K_f}, \tilde{E}_{K_f})$. If K_f' is a compact open subgroup of K_f then we have the pullback map

$$H^*(X_{K_f}, \tilde{E}_{K_f}) \to H^*(X_{K_f'}, \tilde{E}_{K_f'}). \tag{3. 1. 1}$$

We adopt Harder's notation [15] and write

$$H^*(\tilde{X}, \tilde{E}) \stackrel{\text{def}}{=} \varinjlim_{K_f} H^*(X_{K_f}, \tilde{E}_{K_f})$$

even though the symbols $\tilde{X}, \tilde{E}$ are given no independent meaning.

Right translation by $g \in G_3(\mathbf{A}_f)$ induces maps

$$X_{K_f} \xrightarrow{R(g)} X_{g^{-1}K_f g}, \quad G_3(\mathbf{Q}) x Z_\infty K_\infty K_f \longmapsto G_3(\mathbf{Q}) x g Z_\infty K_\infty (g^{-1} K_f g)$$

which in turn induce isomorphisms in cohomology

$$H^*(X_{K_f}, \tilde{E}_{K_f}) \xleftarrow{R(g)^*} H^*(X_{g^{-1}K_f g}, \tilde{E}_{g^{-1}K_f g}) \tag{3.1.2}$$

by pullback. The maps $R(g)$ commute with the inclusions (3.1.1) in the obvious way. Thus the maps (3.1.2) induce an automorphism of $H^*(\tilde{X}, \tilde{E})$. The resulting map $G_3(\mathbf{A}_f) \to \mathrm{Aut}(H^*(\tilde{X}, \tilde{E}))$ is an admissible left action. Moreover, if K_f is a compact open subgroup of $G_3(\mathbf{A}_f)$ then

$$H^*(\tilde{X}, \tilde{E})^{K_f} \cong H^*(X_{K_f}, \tilde{E}_{K_f}). \tag{3.1.3}$$

For a positive integer N we define $K(N) = \{g \in \prod_p G_3(\mathbf{Z}_p) \mid g \equiv 1 \pmod N\}$ and say that a compact open subgroup $K_f \subseteqq G_3(\mathbf{A}_f)$ has level N if N is the least positive integer for which $K(N) \subseteqq K_f$.

For K_f of level N we set

$$S_{K_f} = K_f \cdot \prod_{p \nmid N} (G_3(\mathbf{Q}_p) \cap M_3(\mathbf{Z}_p)). \tag{3.1.4}$$

where $M_3(\mathbf{Z}_p)$ is the set of 3×3 matrices with entries in $\mathbf{Z}_p$. Then (K_f, S_{K_f}) is a Hecke pair and we may form the Hecke algebra $\mathscr{H} = \mathscr{H}(K_f, S_{K_f})$. For each prime p not dividing N let $T_{p,1}, T_{p,2}, T_{p,3} \in \mathscr{H}$ be the elements associated to the following double cosets:

$$\begin{aligned} T_{p,1} &\longleftrightarrow K_f \begin{pmatrix} p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_p K_f, \\ T_{p,2} &\longleftrightarrow K_f \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & 1 \end{pmatrix}_p K_f, \\ T_{p,3} &\longleftrightarrow K_f \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix}_p K_f. \end{aligned} \tag{3.1.5}$$

Then $\mathscr{H}$ is commutative and is generated by the set $\{T_{p,1}, T_{p,2}, T_{p,3} : p \nmid N\}$. As in § 1.1 $\mathscr{H}$ acts on $H^*(\tilde{X}, \tilde{E})^{K_f}$ and thus also on $H^*(X_{K_f}, \tilde{E}_{K_f})$ by transport of structure.

The center $Z(\mathbf{A}_f)$ acts on $H^*(X_{K_f}, \tilde{E}_{K_f})$ as a group of nebentype operators. For $a \in \mathbf{A}_f^*$ we write $[a]_{K_f}$ for the nebentype operator associated to the central element aI. Let $\chi_\infty : Z_\infty \to \mathbf{C}^*$ be the central character of E. Then a simple calculation shows that for $a \in \mathbf{Q}^*$ we have

$$[a_f]_{K_f} = \chi_\infty(a_\infty). \tag{3.1.6}$$

Thus the characters of $Z(\mathbf{A}_f)$ occurring in $H^*(X_{K_f}, \tilde{E}_{K_f})$ are Hecke characters of type χ_∞ and conductor a divisor of N.

We would like to relate this adelic Hecke algebra to the Hecke algebras of sections 1 and 2. To do this it is convenient to impose the following condition on K_f:

$$\det(K_f) = \prod_p \mathbf{Z}_p^*. \tag{3.1.7}$$

Under this hypothesis, the strong approximation theorem for $SL(3)$ implies

$$K_f G_3^+(\mathbf{Q}) = G_3(\mathbf{A}_f)$$

where $G_3^+(\mathbf{Q})$ is the subgroup of $G_3(\mathbf{Q})$ of matrices with positive determinant. In particular if we set

$$\Gamma = K_f \cap G_3^+(\mathbf{Q}), \quad S = S_{K_f} \cap G_3^+(\mathbf{Q}) \tag{3.1.8}$$

then the Hecke pairs $(\Gamma, S) \subseteqq (K_f, S_{K_f})$ are compatible and the natural map

$$\iota : \mathscr{H}(K_f, S_{K_f}) \to \mathscr{H}(\Gamma, S) \tag{3.1.9}$$

is an isomorphism. Moreover the map

$$X_\Gamma \to X_{K_f}, \quad \Gamma g_\infty K_\infty \mapsto G_3(\mathbf{Q})(g_\infty, 1) K_\infty K_f$$

is a homeomorphism, and the induced isomorphism in cohomology

$$H^*(X_\Gamma, \tilde{E}_\Gamma) \to H^*(X_{K_f}, \tilde{E}_{K_f}) \tag{3.1.10}$$

commutes with the action of $\mathscr{H}(K_f, S_{K_f})$.

Of special interest to us will be the groups

$$K_1(N) = \left\{ g \in \prod_p G_3(\mathbf{Z}_p) \;\middle|\; g \equiv \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 1 \end{pmatrix} \pmod N \right\},$$

$$K_0(N) = \left\{ g \in \prod_p G_3(\mathbf{Z}_p) \;\middle|\; g \equiv \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix} \pmod N \right\}, \tag{3.1.11}$$

$$\Gamma_1(3, N) = K_1(N) \cap G_3^+(\mathbf{Q}),$$

$$\Gamma_0(3, N) = K_0(N) \cap G_3^+(\mathbf{Q}).$$

Then $K_1(N)$, $K_0(N)$ satisfy (3. 1. 7) and we can apply the above remarks. The group $\Gamma_0(3, N)$ normalizes $\Gamma_1(3, N)$ so that the quotient group acts on the cohomology of $\Gamma_1(3, N)$ as a group of nebentype operators. For $a \in \prod_p \mathbf{Z}_p^*$ let $[a]$ denote the nebentype operator associated to an element of $\Gamma_0(3, N)$ whose lower right hand corner is congruent to a modulo N.

The following dictionary is easily established.

Proposition 3. 1. 12. *Under the identification*

$$H^*(X_{K_1(N)}, \tilde{E}_{K_1(N)}) = H^*(X_{\Gamma_1(3,N)}, \tilde{E}_{\Gamma_1(3,N)})$$

of (3. 1. 10) *we have the following identification of adelic Hecke operators and* $\Gamma_1(3, N)$ *double coset operators.*

For $p \nmid N$

(i) $$T_{p,1} \longleftrightarrow \Gamma_1(3, N) \begin{pmatrix} p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Gamma_1(3, N),$$

(ii) $$T_{p,2} \longleftrightarrow \Gamma_1(3, N) \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & 1 \end{pmatrix} \Gamma_1(3, N).$$

For $a \in \prod_p \mathbf{Z}_p^*$

(iii) $[a]_{K_f} \longleftrightarrow [a]$.

3. 2 Cohomology at infinity

Let $K_f \subseteqq G_3(\mathbf{A}_f)$ be a compact open subgroup of level N satisfying (3. 1. 7) and let $\mathscr{H} = \mathscr{H}(K_f, S_{K_f})$ be the associated Hecke algebra. Let $\bar{X}_{K_f}$ be the Borel-Serre compactification [4] of X_{K_f}. We will write ∂X_{K_f} for the boundary of $\bar{X}_{K_f}$. If E is a finite dimensional irreducible representation of $G_3(\mathbf{R})$, then the sheaf $\tilde{E}_{K_f}$ over X_{K_f} extends to a sheaf on $\bar{X}_{K_f}$. In this section we examine the systems of Hecke eigenvalues occurring in the cohomology groups

$$H^*(\partial X_{K_f}, \tilde{E}_{K_f}).$$

As in the last section this is closely related to the problem of determining the structure of the $G_3(\mathbf{A}_f)$-module

$$H^*(\partial \tilde{X}, \tilde{E}) \stackrel{\text{def}}{=} \varinjlim_{K_f} H^*(\partial X_{K_f}, \tilde{E}_{K_f}).$$

We will prove the following result.

Proposition 3. 2. 1. *Let* E *be a finite dimensional irreducible complex representation of* $G_3(\mathbf{R})$, *let* $K_f \subseteqq G_3(\mathbf{A}_f)$ *be a compact open subgroup of level* N, *and let*

$$\Phi: \mathscr{H}(K_f) \to \mathbf{C}$$

be a system of eigenvalues occurring in

$$H^r(\partial X_{K_f}, \tilde{E}_{K_f}).$$

Set $b_{p,1}=\Phi(T_{p,1})$, $b_{p,2}=\Phi(T_{p,2})$ and $b_{p,3}=\Phi(T_{p,3})$ for each prime p not dividing N. Let F be a number field which contains all of these eigenvalues and $\mathcal{O}_\lambda$ be the completion of the integers of F at a prime λ with residue characteristic l.

Then there is a **reducible** *Galois representation*

$$\rho_\lambda\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to GL(3, \mathcal{O}_\lambda)$$

unramified outside Nl such that for every $p \nmid Nl$ we have

$$\det(1-\rho_\lambda(\mathrm{Frob}_p)T) = 1 - b_{p,1}T + pb_{p,2}T^2 - p^3 b_{p,3}T^3.$$

The structure of the boundary ∂X_{K_f} is described in [21]. It is a union of three subspaces, W_{P_i,K_f}, $i=0, 1, 2$ associated to the three standard parabolic subgroups

$$P_0=\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix};\quad P_1=\begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix};\quad P_2=\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

The space W_{P,K_f} is homotopically equivalent to the space $P(\mathbf{Q})\backslash G_3(\mathbf{A})/K_\infty K_f$. Moreover, we have

$$W_{P_0,K_f} = W_{P_1,K_f} \cap W_{P_2,K_f}.$$

Thus we may form the Mayer-Vietoris sequence as in [21] but with twisted coefficients.

$$\begin{aligned}(3.2.2)\qquad \cdots \to H^{r-1}(W_{P_0,K_f}, \tilde{E}_{K_f}) &\to H^r(\partial X_{K_f}, \tilde{E}_{K_f}) \\ &\to H^r(W_{P_1,K_f}, \tilde{E}_{K_f}) \oplus H^r(W_{P_2,K_f}, \tilde{E}_{K_f}) \to \cdots\end{aligned}$$

Passing to the inductive limit over K_f we obtain an exact sequence

$$(3.2.3)\qquad \cdots \to H^{r-1}(\tilde{W}_{P_0}, \tilde{E}) \to H^r(\partial\tilde{X}, \tilde{E}) \to H^r(\tilde{W}_{P_1}, \tilde{E}) \oplus H^r(\tilde{W}_{P_2}, \tilde{E}) \to \cdots$$

It is easy to check that the action of $G_3(\mathbf{A}_f)$ commutes with this sequence. It follows that (3.2.2) is a sequence of $\mathscr{H}$-modules.

We introduce a bit of notation. For an algebraic subgroup H of G_3 we will write X_{H,K_f} for the space

$$H(\mathbf{Q})\backslash H(\mathbf{A})/K_\infty^H K_f^H$$

where $K_\infty^H = K_\infty \cap H(\mathbf{R})$ and $K_f^H = K_f \cap H(\mathbf{A}_f)$. We will use the same symbol $\tilde{E}_{K_f}$ for the restriction of $\tilde{E}_{K_f}$ to X_{H,K_f} and set

$$H^r(\tilde{X}_H, \tilde{E}) \overset{\text{def}}{=} \varinjlim_{K_f} H^r(X_{H,K_f}, \tilde{E}_{K_f}).$$

Clearly, $H^r(\tilde{X}_H, \tilde{E})$ is a left $H(\mathbf{A}_f)$-module.

Lemma 3.2.4. *Let P be one of the groups P_0, P_1, P_2. Then there is a natural isomorphism of $G_3(\mathbf{A}_f)$-modules*

$$H^r(\tilde{W}_P, \tilde{E}) \cong \mathrm{Ind}_{P(\mathbf{A}_f)}^{G_3(\mathbf{A}_f)} H^r(\tilde{X}_P, \tilde{E}),$$

where Ind *is defined as in* 1.1 *except that functions are now required to be locally constant.*

Proof. This is a consequence of the structure of the boundary (see [21], esp. 3.6 (2), and compare [15], p. 117). □

Now write

$$P = L \cdot U$$

where L is the Levi component and U is the unipotent radical of P. Let $\mathfrak{u}$ be the Lie algebra of U.

Lemma 3.2.5. *There is a natural isomorphism of $P(\mathbf{A}_f)$-modules*

$$H^r(\tilde{X}_P, \tilde{E}) \cong \bigoplus_{s+t=r} H^s(\tilde{X}_L, \tilde{H}^t(\mathfrak{u}, E)).$$

Remark. Note in particular that the action of $U(\mathbf{A}_f)$ on these groups is trivial.

Proof. For each K_f consider the fibration

$$\begin{array}{ccc} X_{U,K_f} & \longrightarrow & X_{P,K_f} \\ & & \downarrow \\ & & X_{L,K_f}. \end{array}$$

Associated to this fibration there is a spectral sequence

$$H^s(X_{L,K_f}, \tilde{H}^t(X_{U,K_f}, \tilde{E}_{K_f})) \Rightarrow H^{s+t}(X_{P,K_f}, \tilde{E}_{K_f}).$$

By Van Est's theorem [35] we have an isomorphism $H^t(X_{U,K_f}, \tilde{E}) \cong H^t(\mathfrak{u}, E)$. Thus after passing to the inductive limit over K_f our spectral sequence takes the form

$$E_2^{s,t} = H^s(\tilde{X}_L, \tilde{H}^t(\mathfrak{u}, E)) \Rightarrow H^{s+t}(\tilde{X}_P, \tilde{E}).$$

Since $E_2^{s,t} = 0$ for $s > 1$ the spectral sequence degenerates at the E_2 term and the lemma follows. □

Proof of Proposition 3.2.1. If Φ is a system of eigenvalues occurring in $H^r(\partial X_{K_f}, \tilde{E}_{K_f})$ then (3.2.2) implies Φ occurs in $H^*(W_{P_i,K_f}, \tilde{E}_{K_f})$ for some $i = 0$, 1, or 2. Let $P = P_i$ and $P = L \cdot U$ be its Levi decomposition. Lemmas 3.2.4 and 3.2.5 show that there is a choice of s and t such that Φ occurs in

$$(\mathrm{Ind}_{P(\mathbf{A}_f)}^{G_3(\mathbf{A}_f)} H^s(\tilde{X}_L, \tilde{H}^t(\mathfrak{u}, E)))^{K_f}.$$

Consider first the case $P = P_1$. Then $L = L' \times T$ where $L' \cong GL(2)$ and $T \cong GL(1)$. By ([14], Theorem 2. 6. 1) we know that $H^i(\tilde{X}_T, \cdot)$ vanishes for $i > 0$. Thus a simple spectral sequence argument yields an isomorphism

$$H^s(\tilde{X}_L, \tilde{H}^t(\mathscr{U}, E)) \cong H^s(\tilde{X}_{L'}, \tilde{H}^0(\tilde{X}_T, \tilde{H}^t(\mathscr{U}, E))).$$

To simplify the notation we write F for the $L'(\mathbf{R}) \times T(\mathbf{A}_f)$-module $H^0(\tilde{X}_T, \tilde{H}^t(\mathscr{U}, E))$. Then F decomposes into a sum of character spaces under the action of $T(\mathbf{A}_f)$:

$$F = \bigoplus_{\chi} F_\chi$$

where χ runs through Hecke characters of $T(\mathbf{A}_f)$. Thus our system of eigenvalues Φ occurs in one of the spaces

$$(\mathrm{Ind}_{P(\mathbf{A}_f)}^{G_3(\mathbf{A}_f)} H^s(\tilde{X}_{L'}, \tilde{F}_\chi))^{K_f}.$$

Now let φ be a Φ-eigenvector in this space. Write π for the representation of $P(\mathbf{A}_f)$ on $H^s(\tilde{X}_{L'}, \tilde{F}_\chi)$, and π' for the restriction of π to $L'(\mathbf{A}_f)$. Then φ may be viewed as a function

$$\varphi\colon G_3(\mathbf{A}_f)/K_f \to H^s(\tilde{X}_{L'}, \tilde{F}_\chi)$$

satisfying $\varphi(bg) = \pi(b)\,\varphi(g)$ for all b in $P(\mathbf{A}_f)$. Since $P(\mathbf{A}_f) \cdot \prod_p G_3(\mathbf{Z}_p) = G_3(\mathbf{A}_f)$ there is a γ in $\prod_p G_3(\mathbf{Z}_p)$ for which $\varphi(\gamma)$ is nonzero. Fix a prime p not dividing the level of K_f and let T_p denote the Hecke operator associated to the double coset

$$L'(\mathbf{Z}_p) \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}_p L'(\mathbf{Z}_p)$$

in the usual way. We now show that $\varphi_0 = \varphi(\gamma)$ is a T_p-eigenvector and express $b_{p,1}$ in terms of the eigenvalue. In this calculation we will use the fact that φ_0 is $L'(\mathbf{Z}_p)$-invariant and also the fact that φ is invariant under left translation by $U(\mathbf{Q}_p)$.

$$b_{p,1}\varphi_0 = (T_{p,1}\varphi)(\gamma)$$

$$= \sum_{a,b=0}^{p-1} \varphi\left(\gamma \begin{pmatrix} p & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_p\right) + \sum_{c=0}^{p-1} \varphi\left(\gamma \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & c \\ 0 & 0 & 1 \end{pmatrix}_p\right) + \varphi\left(\gamma \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & p \end{pmatrix}_p\right)$$

$$= \sum_{a,b=0}^{p-1} \varphi\left(\begin{pmatrix} p & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_p \gamma\right) + \sum_{c=0}^{p-1} \varphi\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & p & c \\ 0 & 0 & 1 \end{pmatrix}_p \gamma\right) + \varphi\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & p \end{pmatrix}_p \gamma\right)$$

$$= \sum_{a,b=0}^{p-1} \pi'\begin{pmatrix} p & a \\ 0 & 1 \end{pmatrix}_p \varphi_0 + \sum_{c=0}^{p-1} \pi'\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}_p \varphi_0 + \chi_p(p)\varphi_0$$

$$= pT_p\varphi_0 + \chi_p(p)\varphi_0.$$

Thus we see that φ_0 is a T_p-eigenvector. Let a_p be the eigenvalue. A similar calculation can be carried out for the operators $T_{p,2}$, $T_{p,3}$. If we let ψ be the central character of $L'(\mathbf{A}_f)$ acting on φ_0 then the result of these calculations is summarized by

$$b_{p,1} = pa_p + \chi_p(p),$$
$$b_{p,2} = \chi_p(p)a_p + p^2\psi_p(p),$$
$$b_{p,3} = \chi_p(p)\,\psi_p(p).$$

We know from $GL(2)$-theory [8] and from class-field theory respectively, that there are Galois representations

$$\rho_0 \colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to GL(2, \mathcal{O}_\lambda),$$
$$\rho_\chi \colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to GL(1, \mathcal{O}_\lambda)$$

unramified outside Nl and satisfying

$$\det(1 - \rho_0(\mathrm{Frob}_p)\,T) = 1 - a_p T + p\psi_p(p)\,T^2,$$
$$\det(1 - \rho_\chi(\mathrm{Frob}_p)\,T) = 1 - \chi_p(p)\,T$$

for all $p \nmid Nl$. The action of the Galois group on l-power roots of unity gives us a character ω_{cycl} satisfying $\omega_{\mathrm{cycl}}(\mathrm{Frob}_p) = p$. Set $\rho_\lambda = (\rho_0 \otimes \omega_{\mathrm{cycl}}) \oplus \rho_\chi$. A simple calculation shows that ρ_λ satisfies the conclusion of the proposition.

The proofs for the cases $P = P_0, P_2$ are similar. □

3.3 Relative Lie algebra cohomology

Suppose E is a finite dimensional irreducible rational representation over $\mathbf{C}$ of $G_3(\mathbf{R})$. Then for compact open subgroups K_f of $G_3(\mathbf{A}_f)$ we have isomorphisms ([5], VII 2.5)

$$H^*(sl_3, so_3;\, C^\infty(G_3(\mathbf{Q})\backslash G_3(\mathbf{A})/Z_\infty K_f) \otimes E) \cong H^*(X_{K_f}, \tilde{E}_{K_f})$$

where sl_3, so_3 are the Lie algebras of $SL(3, \mathbf{R})$, $SO(3)$ respectively, and C^∞ denotes smooth functions. These isomorphisms commute with the maps (3.1.2). Thus, passing to the inductive limit over K_f we obtain an isomorphism of $G_3(\mathbf{A}_f)$-spaces

$$(3.3.1) \qquad H^*(sl_3, so_3;\, C^\infty(G_3(\mathbf{Q})\backslash G_3(\mathbf{A})/Z_\infty)^o \otimes E) \cong H^*(\tilde{X}, \tilde{E})$$

where $C^\infty(\ )^o$ is the space of $K(1)$-finite smooth functions. We will use this isomorphism in lemma 3.4.3 to compute eigenvalues of Hecke operators.

In 3.5 we will need the following vanishing result. Recall the definition of the interior cohomology $H_!$ from section 1.4.

Lemma 3.3.2. *Suppose E is not isomorphic to its own contragredient. Let Γ be an arithmetic subgroup of $SL(3, \mathbf{Q})$. Then*

$$H_!^*(\Gamma, E) = 0.$$

Proof. Any interior cohomology class $\varphi \in H_!^*(\Gamma, E)$ can be represented by a smooth compactly supported $\tilde{E}_\Gamma$-valued differential form on X_Γ. We choose an admissible scalar product on E as in ([3], § 5.1; [5], II § 2). Then φ can be represented by an L^2 harmonic form ([23], p. 165). Propositions 5.5 and 5.6 of [3] now show that φ is in the image of the canonical map

$$H^*(sl_3, so_3; L^2_{\text{dis}}(\Gamma\backslash SL(3, \mathbf{R}))^\infty \otimes E) \to H^*(\Gamma, E) \tag{3.3.3}$$

where $L^2_{\text{dis}}(\ )^\infty$ is the space of smooth vectors in the discrete spectrum of L^2. The left hand side of (3.3.3) decomposes into a finite direct sum

$$\oplus H^*(sl_3, so_3; H_i^\infty \otimes E) \tag{3.3.4}$$

where for each i, H_i^∞ is the space of smooth vectors in a complete irreducible unitary representation of $SL(3, \mathbf{R})$. But proposition 6.12 II of [5] and our assumption that E is not self dual imply that the space (3.3.4) vanishes. This proves $\varphi = 0$. □

3.4 Symmetric squares

For the rest of the paper g denotes a nonnegative integer and R is a ring in which $g!$ is invertible. If M is a free R-module of finite rank there is a canonical splitting

$$\overset{g}{\otimes} M = \mathrm{Sym}^g(M) \oplus W$$

where $\mathrm{Sym}^g(M)$ is the module of symmetric tensors. The natural isomorphism

$$(\overset{g}{\otimes} M)^* = \overset{g}{\otimes} (M^*)$$

induces an isomorphism $\mathrm{Sym}^g(M)^* = \mathrm{Sym}^g(M^*)$.

Definition 3.4.1. For $n \geq 1$ let M_n be the left $GL(n, R)$-module of column vectors R^n and set $S_g^n(R) = \mathrm{Sym}^g(M_n)$.

If $X_1, \ldots, X_n$ is the standard basis for M_n we may identify $S_g^n(R)$ with the module of degree g polynomials over R in $(X_1, \ldots, X_n)$. The action of $\sigma \in GL(n, R)$ is given by $(\sigma F)(X_1, \ldots, X_n) = F((X_1, \ldots, X_n)\sigma)$. Similarly if $(\xi_1, \ldots, \xi_n)$ is the dual basis in M_n^* to $(X_1, \ldots, X_n)$ then $S_g^n(R)^*$ may be identified with degree g polynomials in $(\xi_1, \ldots, \xi_n)$. The $GL(n, R)$-action is given by

$$(\sigma G)(\xi_1, \ldots, \xi_n) = G((\xi_1, \ldots, \xi_n)\, {}^t\sigma^{-1})$$

where ${}^t\sigma$ is the transpose of σ.

Multiplication by $\Delta^n = \sum_{i=1}^{n} X_i \otimes \xi_i$ induces a $GL(n, R)$-morphism

$$S^n_{g-1}(R) \otimes S^n_{g-1}(R)^* \xrightarrow{\Delta^n_g} S^n_g(R) \otimes S^n_g(R)^*.$$

Definition 3.4.2. (a) $V_g(R)$ is the $GL(2, R)$-module $S^2_g(R)$.

(b) $W_g(R)$ is the $GL(3, R)$-module $\operatorname{coker}(\Delta^3_g)$.

Elements of $W_g(R)$ will be denoted by representatives in $S^n_g(R) \otimes S^n_g(R)^*$ when no confusion will arise.

Our assumption that $g!$ is invertible in R assures that $V_g(R)$ is an irreducible $GL(2, R)$-module and also that $W_g(R)$ is an irreducible $GL(3, R)$-module.

Note that $V_g(\mathbf{C})$ runs through all irreducible rational representations of $SL(2, \mathbf{C})$ and $W_g(\mathbf{C})$ through all self-dual irreducible representations of $SL(3, \mathbf{C})$ as $g = 0, 1, 2, \ldots$.

Recall the definition of $H_!$ from section 1.4 and the definition of $T_{p,1}$, $T_{p,2}$ from the end of 3.1 where we take $N = 1$.

Lemma 3.4.3. *If θ in $H^1_!(SL(2, \mathbf{Z}), V_g(\mathbf{C}))$ is an eigenclass for all the Hecke operators T_p with eigenvalues a_p then there exists Θ in $H^3_!(SL(3, \mathbf{Z}), W_g(\mathbf{C}))$ which is an eigenclass for all the Hecke operators $T_{p,1}$ and $T_{p,2}$ with eigenvalues respectively $p^{-g}(a_p^2 - p^{g+1})$ and $p^{-g}(a_p^2 - p^{g+1})$* (sic).

Remark. Note that "eigenclass" implies by definition that θ and Θ are nonzero.

Proof. Given θ, there exists a holomorphic cusp form in the classical sense of weight $g+2$ for $SL(2, \mathbf{Z})$ with the same Hecke eigenvalues as θ ([32], Chapter 8). Corresponding to the latter we have an irreducible cuspidal automorphic representation $\pi = \otimes \pi_v$ of $GL(2, \mathbf{A})$ with trivial central character ([11], 5.19). The local representation π_∞ is the discrete series representation with lowest weight $g+2$ and trivial central character; for each prime p, π_p is a principal series representation $\pi(\mu_p, \mu_p^{-1})$ where μ_p is an unramified unitary character satisfying $a_p = p^{\frac{g+1}{2}}(\mu_p(p) + \mu_p^{-1}(p))$ ([11], 5.21). By [25] we know that π is not a monomial representation.

Gelbart and Jacquet ([10], theorem 3; [12], theorem 9.3) have shown that given a nonmonomial π as in the last paragraph there exists an irreducible cuspidal automorphic representation $\Pi = \otimes \Pi_v$ of $GL(3, \mathbf{A})$ which is a "symmetric square lift" of π. This means that for each prime p, Π_p is the principal series representation $\Pi(\mu_p^2, 1, \mu_p^{-2})$ ([12], § 3). The representation Π_∞ is described in terms of the associated representation of the Weil group. This can be translated to an explicit description of Π_∞ using theorem 4.4.1 of [18]. In this way one finds that Π_∞ is induced from the standard parabolic subgroup $P \subseteqq GL(3)$ of type (2, 1) as follows. Let $P(\mathbf{R}) = {}^0M \cdot A \cdot N$ be the Langlands decomposition of $P(\mathbf{R})$ ([5], III 3.2). Then ${}^0M$ is isomorphic to

$$SL^{\pm}(2, \mathbf{R}) = \{g \in GL(2, \mathbf{R}) \mid \det(g) = \pm 1\}.$$

Let σ_{2g+3} be the discrete series representation of ${}^0M$ with lowest weight $2g+3$ and χ_0 be the trivial character of AN. Then Π_∞ is the unitarily induced representation $I_{P, \sigma_{2g+3}, \chi_0}$ ([5], III 3.2).

Since Π is a direct summand of the space $L_0^2(G_3(\mathbf{Q})\backslash G_3(\mathbf{A})/Z_\infty)^\infty$ of smooth L^2 cuspidal functions we have an inclusion of $G_3(\mathbf{A}_f)$-modules

$$H^3(sl_3, so_3; \Pi \otimes W_g(\mathbf{C})) \subseteq H^3(sl_3, so_3; L_0^2(\ )^\infty \otimes W_g(\mathbf{C})).$$

Taking $K(1)$-invariants we obtain an inclusion of $\mathscr{H}(K(1), S_{K(1)})$-modules

$$(3.4.4) \qquad H^3(sl_3, so_3; \Pi_\infty \otimes W_g(\mathbf{C})) \otimes \bigotimes_p \Pi_p^{G_3(\mathbf{Z}_p)} \hookrightarrow H^3(sl_3, so_3; L_0^2(G_3(\mathbf{Q})\backslash G_3(\mathbf{A})/Z_\infty K(1))^\infty \otimes W_g(\mathbf{C})).$$

By ([2], cor. 5.5) the natural map from this latter group to $H^3(SL(3, \mathbf{Z}, W_g(\mathbf{C}))$ is injective and has image contained in $H_!^3(SL(3, \mathbf{Z}), W_g(\mathbf{C}))$.

On the other hand a calculation based on ([5], III 3.3) proves

$$H^3(sl_3, so_3; \Pi_\infty \otimes W_g(\mathbf{C})) \cong \mathbf{C}.$$

Clearly $\Pi_p^{G_3(\mathbf{Z}_p)}$ is one dimensional for each prime p. Hence the left hand side of (3.4.4) is isomorphic to $\mathbf{C}$. The eigenvalues of $T_{p,1}$ and $T_{p,2}$ acting on this space are easily calculated and seen to be those given in the statement of the proposition. □

3.5 Torsion

If l is a rational prime, the Teichmüller character is the unique character $\omega: \mathbf{F}_l^* \to \mathbf{Z}_l^*$ satisfying the congruence $\omega(a) \equiv a \pmod{l}$. Evaluation of ω on the lower right entry of a matrix in $\Gamma_0(3, l)$ induces a character $\Gamma_0(3, l)/\Gamma_1(3, l) \to \mathbf{Z}_l^*$ which we will also denote by ω.

Proposition 3.5.1. *Let θ be as in Lemma 3.4.3, $l > g$ a rational prime, $l \neq 2, 3$, and K a finite extension of $\mathbf{Q}$. Let λ be a prime of K lying over l and $\mathcal{O}_\lambda$ the ring of integers in K_λ. For any $p \neq l$, let $T_{p,1}$ and $T_{p,2}$ denote the Hecke operators associated to $\Gamma_1(3, l)$ as in 3.1. Set $b_p = p^{-g}(a_p^2 - p^{g+1})$.*

(1) *Then for a suitable choice of K there is a nonzero*

$$\Theta^0 \in H^3(\Gamma_1(3, l), S_g^3(\mathcal{O}_\lambda))(\omega^g)$$

such that Θ^0 is an eigenvector for all $T_{p,1}$ and $T_{p,2}$, $p \neq l$, with eigenvalues $b_{p,1}$ and $b_{p,2}$ respectively, satisfying the congruence $b_{p,1} \equiv b_{p,2} \equiv b_p \pmod{\lambda}$.

(2) *If we fix g and θ, then for sufficiently large l the class Θ^0 is a torsion class.*

Proof. We make the following assignments:

$$\Gamma = SL(3, \mathbf{Z}), \quad \Gamma_0 = \Gamma_0(3, l),$$
$$S = S_{K(1)} \cap G_3^+(\mathbf{Q}),$$
$$S_0 = S_{K_0(l)} \cap G_3^+(\mathbf{Q}).$$

As in Lemma 1. 2. 4, we know that there is a finite extension K of $\mathbf{Q}$ which will contain all the Hecke-eigenvalues appearing in cohomology groups we need to deal with. We fix such a K and set $R=\mathcal{O}_\lambda$, $P=(\lambda)$.

By lemma 3. 4. 3 there is a class Θ in $H^3(SL(3,\mathbf{Z}), W_g(R))$ such that for every prime $p\neq l$, Θ is an eigenvector for $T_{p,1}$ and $T_{p,2}$ with eigenvalues $b_p=p^{-g}(a_p^2-p^{g+1})$.

We define $\psi\colon S_g^3(\bar{R})\otimes R_{\omega^{-g}}\to W_g(\bar{R})$ by $\psi(F\otimes 1)=F\otimes\xi_3^g$ (notation of 3. 4). This is an S_0-morphism.

As in 1. 3 we have an RS-morphism $\beta(\psi)\colon \mathrm{Ind}(\Gamma_0,\Gamma,S_g^3(\bar{R})\otimes\bar{R}_{\omega^{-g}})\to W_g(\bar{R})$ and we may draw diagram 1. 3. 3. Clearly, $X_1^g\otimes\xi_3^g$ is a nonzero element of the image of $\beta(\psi)$. Since $\bar{E}$ is irreducible we must have $\beta(\psi)$ is surjective. By theorem 1. 3. 5, B^N is surjective. Assertion (1) follows from theorem 1. 3. 4 (a) and lemma 1. 1. 5.

A theorem of Deligne [8] proves that there is a two dimensional irreducible λ-adic representation σ_λ of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ which is unramified outside l and such that the characteristic polynomial of $\sigma_\lambda(\mathrm{Frob}_p)$ is $1-a_pT+pT^2$. Let σ_λ^2 be the symmetric square of this representation and let ω_{cycl} be the character of the Galois group acting on l-power roots of unity. Then one easily verifies the identity

$$\det\left(1-(\sigma_\lambda^2\otimes\omega^{-g})(\mathrm{Frob}_p)\right)=1-b_pT+pb_pT^2-p^3T^3.$$

By a theorem of Ribet [25], $\sigma_\lambda(\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}))\bmod\lambda$ contains $SL(2,\mathcal{O}/\lambda)$ for almost all l. For such l, $\sigma_\lambda^2(\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}))\bmod\lambda$ contains $SL(3,\mathcal{O}/\lambda)$. Hence there is a $B>0$ such that whenever $l>B$, $\sigma_\lambda^2\bmod\lambda$ is irreducible.

Now fix $l>B$. We claim that Θ^0 is a torsion class. For suppose Θ^0 were not torsion. Then by lemma 3. 3. 2 we know that the restriction of Θ^0 to the boundary does not vanish. Let $b_{p,i}$ be the eigenvalues of $T_{p,i}$, $i=1,2,3$, acting on Θ^0. Then theorem 3. 2. 1 says that there is a reducible three dimensional λ-adic Galois representation ρ_λ such that

$$\det\left(1-\rho_\lambda(\mathrm{Frob}_p)T\right)=1-b_{p,1}T+pb_{p,2}T^2-p^3b_{p,3}T^3.$$

By (1) we have congruences $b_{p,1}\equiv b_{p,2}\equiv b_p\ (\mathrm{mod}\,\lambda)$. Using (3. 1. 12 (iii)) and (3. 1. 6) we find $b_{p,3}=p^g\omega(p)^{-g}$. In particular $b_{p,3}\equiv 1\ (\mathrm{mod}\,\lambda)$. Thus we have a congruence

$$\det\left(1-\rho_\lambda(\mathrm{Frob}_p)T\right)\equiv\det\left(1-(\sigma_\lambda^2\otimes\omega^{-g})(\mathrm{Frob}_p)T\right)\ (\mathrm{mod}\,\lambda)$$

for every prime $p\neq l$. But ρ_λ is reducible and $\sigma_\lambda^2\otimes\omega^{-g}$ is irreducible, so by the Čebotarev density theorem this contradicts the Brauer-Nesbitt theorem. □

Lemma 3. 5. 2. *Let f and f' be two classical holomorphic cusp forms of the same weight k for the full modular group $SL(2,\mathbf{Z})$. Assume each is an eigenform for all the Hecke operators T_p with eigenvalues a_p and a'_p respectively, for all primes p. Suppose $a_p=\pm a'_p$ for every p. Then f and f' are proportional.*

Proof. We prove the lemma by using the principle of [25] that the l-adic representations of f and f' are "as independent as possible."

Let $E=\mathbf{Q}[a_p|p\text{ prime}]=\mathbf{Q}[a'_p|p\text{ prime}]$ and $\mathcal{O}$ be the ring of integers of E. For a prime l let $\mathcal{O}_l=\mathcal{O}\otimes_{\mathbf{Z}}\mathbf{Z}_l$.

We first show that there is a $\sigma \in \mathrm{Aut}(\mathbf{C})$ such that $\sigma(a_p) = a'_p$ for all primes p. For suppose there is no such σ. Let $(\alpha, \beta) \in \mathcal{O} \times \mathcal{O}$ such that $E \times E = \mathbf{Q}[(\alpha^2, \beta^2)]$. For sufficiently large l we have $\mathbf{Z}_l[(\alpha^2, \beta^2)] = \mathcal{O}_l \times \mathcal{O}_l$ and by theorem 6.1 of [25] there is a prime p such that $(a_p, a'_p) \equiv (\alpha, \beta) \pmod{l}$. Then $(a_p^2, a_p'^2)$ generates $E \times E$ over $\mathbf{Q}$. This contradicts our hypothesis $a_p = \pm a'_p$.

By lemma 4.8 of [25] there is a prime p such that $\mathbf{Q}[a_p^2] = E$. Since

$$\sigma(a_p) = a'_p = \pm a_p$$

we have $\sigma | E$ is the identity map on E. In particular $\sigma(a_p) = a_p = a'_p$, proving the lemma. □

Theorem 3.5.3. *Let g be a fixed positive integer and set $d(g) =$ the dimension of the space of holomorphic cusp forms for $SL(2, \mathbf{Z})$ of weight $g+2$. Then if l is sufficiently large,*

$$\dim_{\mathbf{F}_l} H^3(\Gamma_1(3, l), S_g^3(\mathbf{Z}))_{l\text{-torsion}} \geqq d(g).$$

Proof. The space of cusp forms mentioned above is naturally a sub-Hecke module of $H_!^1(SL(2, \mathbf{Z}), V_g(\mathbf{C}))$ by Eichler-Shimura. Set $d = d(g)$. By lemmas 3.4.3 and 3.5.2 there are d linearly independent Hecke eigenclasses $\Theta_1, \ldots, \Theta_d$ in $H_!^3(SL(3, \mathbf{Z}), W_g(\mathbf{C}))$ with eigenvalues $p^{-g}(a_p(i)^2 - p^{g+1})$, $i = 1, \ldots, d$.

Let l be greater than g, large enough so that the d infinite-dimensional vectors

$$(\ldots, \bar{p}^{-g}(\bar{a}_p(i)^2 - \bar{p}^{g+1}), \ldots)$$

are distinct (bar denotes reduction mod λ), and large enough so that the conclusion of (2) of Proposition 3.5.1 is valid. Then corresponding to these systems of eigenvalues there are d linearly independent eigenclasses $\Theta_1^0, \ldots, \Theta_d^0$ in

$$H^3(\Gamma_1(3, l), S_g^3(\mathcal{O}_\lambda))_{\lambda\text{-torsion}}.$$

The theorem now follows immediately. □

Remark. In particular, the dimension of the l-torsion in $H^3(\Gamma_1(3, l), S_g^3(\mathbf{Z}))$ becomes arbitrarily large as $g \to \infty$ and l is sufficiently large. Compare this with the fact that there is *no* l-torsion in $H^1(\Gamma_1(2, l), S_g^2(\mathbf{Z}))$ whenever $l > g$.

References

[1] *A. N. Andrianov*, The multiplicative arithmetic of Siegel modular forms, Russian Math. Surveys **34** (1979), 75—148.

[2] *A. Borel*, Stable real cohomology of arithmetic groups. II, in: Manifolds and Lie Groups, Progress in Mathematics **14**, Boston 1980.

[3] *A. Borel, H. Garland*, Laplacian and the discrete spectrum of an arithmetic group, Amer. J. Math. **105** (1983), 309—335.

[4] *A. Borel, J.-P. Serre*, Corners and arithmetic groups, Comm. Math. Helv. **48** (1973), 436—491.

[5] *A. Borel, N. Wallach*, Continuous cohomology, discrete subgroups, and representations of reductive groups, Ann. Math. Studies **94**, Princeton 1980.

[6] *K. S. Brown*, Cohomology of infinite groups, Proc. Int. Cong. Math. I (1978), 285—290.

[7] *H. Cartan*, Cohomologie des groupes, suites spectrales, faisceaux, Seminaire ENS 1950/51, New York 1967.

[8] *P. Deligne*, Formes modulaires et représentations l-adiques, Sem. Bourbaki, Lect. Notes in Math. **179** (1971), 136—186.
[9] *P. Deligne, J.-P. Serre*, Formes modulaire de poids 1, Ann. scient. Ec. Norm. Sup. (4) **7** (1974), 507—530.
[10] *S. Gelbart*, Automorphic forms and Artin's conjecture, in: Modular Functions of One Variable VI, Lecture Notes in Math. **627** (1976), 241—276.
[11] *S. Gelbart*, Automorphic forms on adele groups, Ann. Math. Studies **83**, Princeton 1975.
[12] *H. Gelbart, H. Jacquet*, A relation between automorphic representations of $GL(2)$ and $GL(3)$, Ann. Scient. Ec. Norm. Sup. (4) **11** (1978), 471—542.
[13] *K. Haberland*, Perioden von Modulformen einer Variablen und Gruppencohomologie. I, II, III, Math. Nachr. **112** (1983), 245—315.
[14] *G. Harder*, Eisenstein cohomology of arithmetic groups: the case GL_2, Preprint 1985.
[15] *G. Harder*, Period integrals of Eisenstein cohomology classes and special values of some L-functions, in: Number theory related to Fermat's Last Theorem, Progress in Mathematics **26**, Boston 1983.
[16] *H. Hida*, Kummer's criterion for the special values of Hecke L-functions of imaginary quadratic fields and congruences among cusp forms, Invent. Math. **66** (1982), 415—459.
[17] *H. Hida*, On congruence divisors of cusp forms as factors of the special values of their zeta-functions, Invent. Math. **64** (1981), 221—262.
[18] *H. Jacquet*, Principal L-functions of the linear group, in: Automorphic Forms, Representations, and L-functions. 2, Proceedings of Symposia in Pure Mathematics **33**, Providence 1979.
[19] *N. Jochnowitz*, A study of the local components of the Hecke algebra mod l, Trans. AMS **270** (1982), 253—267.
[20] *M. Kuga, W. Parry, C.-H. Sah*, Group cohomology and Hecke operators, in: Manifolds and Lie Groups, Progress in Mathematics **14**, Boston 1980.
[21] *R. Lee, J. Schwermer*, Cohomology of arithmetic subgroups of $SL(3)$ at infinity, J. reine angew. Math. **330** (1982), 100—131.
[22] *B. Mazur, A. Wiles*, Class fields of abelian extensions of **Q**, Invent. Math. **76** (1984), 179—330.
[23] *Georges de Rham*, Variétés différentiables. Paris 1955.
[24] *K. Ribet*, Mod p Hecke operators and congruences between modular forms, Invent. Math. **71** (1983), 193—205.
[25] *K. Ribet*, On l-adic representations attached to modular forms, Invent. Math. **28** (1975), 245 275.
[26] *J.-P. Serre*, Cohomologie des groupes discrets, Ann. Math. Studies **70** (1971), 77—169.
[27] *J.-P. Serre*, Corps Locaux, Publications de l'Institut de Mathématique de l'Université de Nancago **8**, Paris 1968.
[28] *J.-P. Serre*, Formes modulaires et fonctions zeta p-adique, Lect. Notes in Math. **350** (1973), 191—269.
[29] *J.-P. Serre*, Letter to J.-M. Fontaine (1979).
[30] *J.-P. Serre*, Linear representations of finite groups, Graduate Texts in Mathematics **42**, Berlin-Heidelberg-New York 1977.
[31] *J.-P. Serre*, Valeurs propres des operateurs de Hecke modulo l, Asterisque **24/25** (1977), 109—117.
[32] *G. Shimura*, Introduction to the arithmetic theory of automorphic forms. Publ. Math. Soc. Japan **11**, Princeton 1971.
[33] *G. Shimura*, An l-adic method in the theory of automorphic forms, Unpublished (1968).
[34] *H. P. F. Swinnerton-Dyer*, On l-adic representations and congruences for coefficients of modular forms. I, II, Lect. Notes in Math. **350** (1973), 1—55; **601** (1977), 63—90.
[35] *Van Est*, A generalization of the Cartan-Leray spectral sequence. II, Indagationes Math. A **20** (1958), 406—413.
[36] *H. Weyl*, The classical groups; their invariants and representations, Princeton Mathematical Series **1**, Princeton 1946.

Department of Mathematics, Ohio State University, Columbus, OH 43210, U.S.A.

Department of Mathematics, Boston University, Boston, MA 02215, U.S.A.

Eingegangen 27. August 1984, in revidierter Form 29. Juni 1985