ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO MATHEMATICA

TOMUS XLVII.

2004

ANNALES

UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO BIOLOGICA
incepit anno MCMLVII
SECTIO CHIMICA
incepit anno MCMLIX
SECTIO CLASSICA
incepit anno MCMXXIV
SECTIO COMPUTATORICA
incepit anno MCMLXXVIII
SECTIO GEOGRAPHICA
incepit anno MCMLXVI
SECTIO GEOLOGICA
incepit anno MCMLVII
SECTIO GEOPHYSICA ET METEOROLOGICA
incepit anno MCMLXXV
SECTIO HISTORICA
incepit anno MCMLVII
SECTIO IURIDICA
incepit anno MCMLIX
SECTIO LINGUISTICA
incepit anno MCMLXX
SECTIO MATHEMATICA
incepit anno MCMLVIII
SECTIO PAEDAGOGICA ET PSYCHOLOGICA
incepit anno MCMLXX
SECTIO PHILOLOGICA
incepit anno MCMLVII
SECTIO PHILOLOGICA HUNGARICA
incepit anno MCMLXX
SECTIO PHILOLOGICA MODERNA
incepit anno MCMLXX
SECTIO PHILOSOPHICA ET SOCIOLOGICA
incepit anno MCMLXII

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 3–16

ON SETS OF UNIQUENESS FOR ADDITIVE AND MULTIPLICATIVE FUNCTIONS OVER THE MULTIPLICATIVE GROUP GENERATED BY THE POLYNOMIAL $x^2 + a$

By

J. FEHÉR* and I. KÁTAI*

(Received April 23, 2003)

Dedicated to the memory of our friend Peter Kiss

§1. Let $\mathbb{N}, \mathbb{Z}, Q$ be the set of natural numbers, integers, rational numbers, respectively. Let $Q_\times$ be the multiplicative group of positive rationals.

For some subset A of $Q_\times$, let $Q_\times(A)$ be the multiplicative group generated by A, that is $\alpha \in Q_\times(A)$ if there exist suitable elements $a_1, \dots, a_r \in A$, and $\varepsilon_1, \dots, \varepsilon_r \in \{-1, 1\}$, such that $\alpha = a_1^{\varepsilon_1} \dots a_r^{\varepsilon_r}$. Let furthermore $B_x(A)$ be the group of those α for which there exist suitable $a_1, \dots, a_r \in A$, $l_1, \dots, l_r \in \mathbb{Z}$, $d \in \mathbb{N}$ such that

$$\alpha^d = a_1^{l_1} \dots a_r^{l_r}.$$

Let $\mathcal{P}$ be the whole set of primes.

Let $a \in \mathbb{N}$, $\mathcal{I}_a = \{n^2 + a \mid n \in \mathbb{N}\}$, $\mathcal{E}_a$ be the set of those primes p, for which $p \mid n^2 + a$ holds for some n. Let $\mathcal{E}_a = \mathcal{E}_a^{(1)} \cup \mathcal{E}_a^{(2)}$, where

$$\mathcal{E}_a^{(1)} = \left\{ p \,\middle|\, \left(\frac{-a}{p} \right) = 1,\ p \in \mathcal{P} \right\},$$

$$\mathcal{E}_a^{(2)} = \{ p \mid p | a,\ p \in \mathcal{P} \}.$$

* Research partially supported by Hungarian National Foundation for Scientific Research Grant and by the research group Applied Number Theory of the Hungarian Academy of Sciences.

Let $\mathcal{A}_a^*$ be the set of those real-valued arithmetical functions f, which are defined originally on $\mathcal{E}_a$, and then extended to $Q_\times(\mathcal{E}_a)$, so that

$$f(rs)=f(r)+f(s), \quad f\left(\frac{r}{s}\right)=f(r)-f(s), \tag{1.1}$$

for every couple of $r, s \in Q_\times(\mathcal{E}_a)$. The extension of the definition of f, by (1.1) is unique.

Let $\tilde{\mathcal{M}}_a^*$ be the set of those g having complex values on the unit circle, defined originally on $\mathcal{E}_a$, and extended to $Q_\times(\mathcal{E}_a)$ such that

$$g(rs)=g(r)g(s), \quad g\left(\frac{r}{s}\right)=g(r)\cdot\overline{g}(s). \tag{1.2}$$

§2. We shall prove

THEOREM 1. *If $f \in \mathcal{A}_a^*$ and $f(n^2+a) \to 0$ as $n \to \infty$, then $f(n^2+a)=0$ for every $n \in \mathbb{N}$.*

If $g \in \mathcal{M}_a^*$ and $g(n^2+a) \to 1$ as $n \to \infty$, then $g(n^2+a)=1 \quad (n \in \mathbb{N})$.

The proof is based upon the following lemmas:

Let $F(n) := n^2+a$.

LEMMA 1. *We have*

$$F(n+F(n)) = F(n)F(n+1). \tag{2.1}$$

PROOF. Obvious. ∎

LEMMA 2. *Assume that $1 < M \in \mathbb{N}$, M is not a square and the equation*

$$u^2+a = M(v^2+a) \tag{2.2}$$

has at least one solution in $u, v \in \mathbb{N}$. Then there is a sequence of couples of positive integers x_ν, y_ν such that $x_\nu, y_\nu \to \infty$ and

$$x_\nu^2+a = M(y_\nu^2+a). \tag{2.3}$$

PROOF. Let us write (1.2) in the form $u^2 - Mv^2 = (M-1)a$. The Pell-equation $u^2 - Mv^2 = 1$ has infinitely many positive solution, (ξ_ν, η_ν). Then the pair, x_ν, y_ν counted from

$$x_\nu + \sqrt{M}y_\nu = \left(\xi_\nu + \sqrt{M}y_\nu\right)\left(u+\sqrt{M}v\right)$$

satisfies (2.3).

The proof of Lemma 2 is finished. ∎

LEMMA 3. *The values $F(n)$, $F(n+1)$ cannot be squares simultaneously.*

PROOF. Assume indirectly that $n^2 + a = U^2$, $(n+1)^2 + a = V^2$. Then $V \geq U + 1$. Furthermore, $2n + 1 = V^2 - U^2 \geq V + U + 1$, whence $n > U$. But it is impossible. ∎

LEMMA 4. *If $F(n + F(n))$ is a square, then none of $F(n)$, $F(n+1)$ can be squares.*

PROOF. Obvious. If either $F(n)$, or $F(n+1)$ would be a square, then the other would be a square as well, which contradicts to Lemma 3. ∎

PROOF OF THEOREM 1. We shall prove the first assertion. The second assertion can be proved completely on the same way.

We can apply Lemma 2 for $M = F(n)$ if M is not a square. We obtain that $f(n^2+a) = 0$ if $n^2+a \neq$ square. Assume that $F(n) =$ square. Since $F(n+1)$, $F(n + F(n))$ are not squares, therefore $f(F(n+1)) = 0$, $f(F(n + F(N))) = 0$, from (1.1) we obtain that $f(F(n)) = 0$.

The proof of Theorem 1 is completed. ∎

§3. A subset $E(\subseteq Q_\times(\mathcal{E}_a))$ is called a set of uniqueness (for the class of functions f in $\mathcal{A}_a^*$), if $f \in \mathcal{A}_a^*$, $f(E) = 0$ implies that $f(Q_\times(\mathcal{E}_a)) = 0$.

Similarly, a subset $F \subseteq Q_\times(\mathcal{E}_a)$ is called a set of uniqueness mod 1 (for the class of functions g in $\mathcal{M}_a^*$), if $g \in \mathcal{M}_a^*$, $g(F) = 1$ implies that $g(Q_\times(\mathcal{E}_a)) = 1$.

It is known that $\mathcal{E}_a$ is a set of uniqueness mod 1, if and only if

$$Q_\times(\mathcal{J}_a) = Q_\times(\mathcal{E}_a), \tag{3.1}$$

and that $\mathcal{E}_a$ is a set of uniqueness, if

$$B_\times(\mathcal{J}_a) = Q_\times(\mathcal{E}_a). \tag{3.2}$$

The notion of "set of uniqueness" for completely additive functions was introduced by Kátai [1]. The assertion formulated in (3.2) is proved by D. Wolke. Relation (3.1) was proved in [4], [5], [7].

We can suggest to read the relevant further papers [8], [9], [10], [11].

§4. The assertion, formulated in the next lemma, is clear.

LEMMA 5. *Let $p \in \mathcal{P}$, $p > \dfrac{2\sqrt{a}}{\sqrt{3}}$, $\left(\dfrac{-a}{p}\right) = 1$. Let x_0 be the smallest positive integer, for which $p \mid x^2 + a$. Let $x_0^2 + a = K_p \cdot p$. Then $K_p = \dfrac{x_0^2 + a}{p} < p$.*

PROOF. We have $x_0 \le p/2$. The assertion is obvious. ∎

LEMMA 6. *Let $a = b^2 c$. Then $a, b^2, c \in Q_\times(\mathcal{I}_a)$.*

PROOF. Let $\varphi(x) = x^2 + a$. Then $\varphi(1) = 1 + a$, $\varphi(a) = a^2 + a$, whence $a = \dfrac{\varphi(a)}{\varphi(1)} \in Q_\times(\mathcal{I}_a)$. Since $\varphi(b) = b^2(1+c)$, $\varphi(bc) = b^2c^2 + b^2c = b^2c(1+c)$, we have that $c = \dfrac{\varphi(bc)}{\varphi(b)} \in Q_\times(\mathcal{I}_a)$. Thus $b^2 = \dfrac{a}{c} \in Q_\times(\mathcal{I}_a)$. ∎

THEOREM 2.

1. *We have $Q_\times(\mathcal{I}_1) = Q_\times(\mathcal{E}_1)$, i.e. $\mathcal{I}_1$ is a set of uniqueness* mod 1.
2. *Let $a = b^2 = p_1^{2\alpha_1} \dots p_r^{2\alpha_r} \cdot q_1^{2\beta_1} \dots q_s^{2\beta_s}$, where*

$$\left(\frac{-1}{p_j}\right) = 1 \qquad (j = 1, \dots, r);$$
$$\left(\frac{-1}{q_l}\right) = -1 \qquad (l = 1, \dots, s).$$

Then

$$Q_\times(\mathcal{I}_a) = Q_\times(\mathcal{I}_1) \times H_a, \tag{4.1}$$

where H_a is the group generated by $\{q_1^2, \dots, q_s^2\}$.

PROOF. To prove the first assertion, we observe, that $1^2 + 1 \in Q_\times(\mathcal{I}_1)$, $2^2 + 1 \in Q_\times(\mathcal{I}_1)$ and by Lemma 5, if $p \in \mathcal{E}_a^{(1)}$, $p \ge 2$, by using induction, then $K_p \in Q_\times(\mathcal{I}_1)$ can be assumed. This completes the proof.

Let us consider now the assertion 2. We have $1 + a \in Q_\times(\mathcal{I}_a)$, $a^2 + a = a(a+1) \in Q_\times(\mathcal{I}_a)$, whence $a \in Q_\times(\mathcal{I}_a)$. Observe furthermore that $b^2 \mathcal{I}_1 \subseteq \mathcal{I}_a$. If a prime π has a representation

$$\pi = \prod \left(n_j^2 + 1\right)^{\varepsilon_i},$$

then

$$\pi = a^{-\sum \varepsilon_i} \prod \left((bn_j)^2 + a \right)^{\varepsilon_i},$$

which by $a \in Q_\times(\mathcal{I}_a)$ implies that $\pi \in Q_\times(\mathcal{I}_a)$.

If $q|a$, $q \equiv -1 \pmod 4$, then $q^\nu \| n^2+a$ implies that ν is an even number. Let $n = q$. Then $q^2 + a = q^2 \left(1 + \left(\frac{b}{q}\right)^2\right)$. We have $q^2 + a \in Q_\times(\mathcal{I}_a)$, and $1 + \left(\frac{b}{q}\right)^2 \subseteq Q_\times(\mathcal{I}_1) \subseteq Q_\times(\mathcal{I}_a)$, thus $q^2 \in Q_\times(\mathcal{I}_a)$.

Consequently, we obtain the second assertion immediately. The theorem is proved. ∎

§5. Assume that

$$a = \pi_1^{2\gamma_1} \dots \pi_r^{2\gamma_r} \quad \rho_1^{2\delta_1+1} \dots \rho_s^{2\delta_s+1}, \tag{5.1}$$

where $\pi_1, \dots, \pi_r, \rho_1, \dots, \rho_s \in \mathcal{P}$, $\gamma_\nu \in \mathbb{N}$, $\delta_\nu \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

By choosing first $b = \pi_j$, $c = \frac{a}{b^2}$, from Lemma 6, we deduce that $\pi_j^2 \in Q_\times(\mathcal{E}_a)$ $(j = 1, \dots, r)$. Similarly, we can choose $b = \rho_j$ if $\delta_j \geq 1$, and deduce that $\rho_j^2 \in Q_\times(\mathcal{E}_a)$. Hence, the next assertion it follows immediately.

THEOREM 3. *Let a as in (5.1). Let H be the group generated by $\pi_1^2, \dots$ $\dots, \pi_r^2$ and those ρ_j^2, for which $\delta_j \geq 1$.*

Let $C = \rho_1 \dots \rho_s$. Then

$$Q_\times(\mathcal{I}_a) \supseteq H \times Q_\times(\mathcal{I}_C). \tag{5.2}$$

§6.

THEOREM 4. *We have*

$$O\left(Q_\times(\mathcal{E}_5) \,/\, Q_\times(\mathcal{I}_5)\right) = 2.$$

The number 2 does not belong to $Q_\times(\mathcal{I}_5)$.

PROOF. Let $\varphi(n) = n^2+5$. If $2|\varphi(n)$, then n is an odd number, $n = 2k+1$, and $\varphi(n) = 2(3 + 2k(k+1))$, i.e. $\frac{\varphi(n)}{2} \equiv -1 \pmod 4$. If n is even, then $\varphi(n) \equiv 1 \pmod 4$.

Assume in contrary that $2 \in Q_\times(\mathcal{I}_5)$. Then

$$2 = \frac{\varphi(n_1)\dots\varphi(n_r)}{\varphi(N_1)\dots\varphi(N_s)} \cdot \frac{\varphi(u_1)\dots\varphi(u_t)}{\varphi(v_1)\dots\varphi(v_h)} = \frac{A}{B}\cdot\frac{C}{D}, \tag{6.1}$$

say, where n_i, N_i are odd, u_l, v_l even numbers.

Thus

$$2BD = AC,$$

$$2^{s+1}\cdot\frac{B}{2^s}D = 2^r\frac{A}{2^r}\cdot C.$$

Since $\frac{B}{2^s}, D, \frac{A}{2^s}, C$ are odd numbers, therefore $r = s+1$. Furthermore $D, C \equiv 1 \pmod 4$ $\frac{B}{2^s} \equiv (-1)^s \pmod 4$, $\frac{A}{2^r} \equiv (-1)^r \pmod 4$, this is impossible.

To finish the proof, we observe that $\varphi(1) = 2\cdot 3$. Let us consider the group $\mathcal{H}$ generated by the union of $\mathcal{I}_5$ and $\{2\}$. Then $2 \in \mathcal{H}$, $3 \in \mathcal{H}$. By using Lemma 5, $\left(5 > \frac{2\sqrt{5}}{3}\right)$, we obtain that $\mathcal{H} = Q_\times(\mathcal{E}_5)$.

The proof is completed. ∎

By the method used above we can prove the following

LEMMA 7. *Assume that* $a \equiv 5 \pmod 8$. *Then* $2 \notin Q_\times(\mathcal{I}_a)$.

The following remark, which we formulate now as Lemma 8, is obvious.

LEMMA 8. *If* $2^2 \| a+1$, *then* $2 \notin Q_\times(\mathcal{I}_a)$.

PROOF. This is clear. If $2 | x^2 + a$, then $x = 2k+1$, and

$$2^2 \| (a+1) + 4k(k+1) = 2^2 \cdot \left\{\frac{a+1}{4} + k(k+1)\right\}.$$

Thus in each expression $\prod_{i=1}^{h}(n_i^2 + a)^{\varepsilon_i}$ the exponent of factor 2 is an even number, consequently 2 cannot be represented in this form. ∎

LEMMA 9. *Let $a = bq$, q a prime, such that* $\left(\frac{b}{q}\right) = 1$. *Then* $q \notin Q_\times(\mathcal{J}_a)$.

PROOF. Let $\varphi(n) = n^2 + a$. Let us observe that $\dfrac{\varphi(qm)}{q} = qm^2 + b$, consequently the left hand side is a quadratic nonresidue mod q. Furthermore, if $q \nmid n$, then $\varphi(n) \pmod q \equiv n^2$, thus it is a quadratic residue mod q.

Assume indirectly that $q \in Q_\times(\mathcal{J}_a)$. Then $\varphi(u_1)\dots\varphi(u_r)\varphi(qn_1)\dots\varphi(qn_s) = q\varphi(v_1)\dots\varphi(v_t)\varphi(qm_1)\dots\varphi(qm_k)$, where $(u_1\dots u_r\ v_1\dots v_t, q) = 1$. Since $q \| \varphi(qm_j)$, $q \| \varphi(qn_l)$, and $(\varphi(u_h), q) = 1$, $(\varphi(v_h), q) = 1$, we obtain that $s = k+1$, and that

$$\varphi(u_1)\dots\varphi(u_r)\, y_1\dots y_s = z_1\dots z_k\ \varphi(v_1)\dots\varphi(v_t), \tag{6.2}$$

where $y_h = \dfrac{\varphi(qn_h)}{q}$, $z_j = \dfrac{\varphi(qm_j)}{q}$.

The value of the Legendre symbol $\left(\frac{\cdot}{q}\right)$ for the integer standing on the left is $(-1)^s$, while the same for the right hand side is $(-1)^k$. Thus $(-1)^s = (-1)^k$, which is impossible. ∎

§7.

THEOREM 5. *For every $a \in \mathbb{N}$, $Q_\times(\mathcal{E}_a)/Q_\times(\mathcal{J}_a)$ is a finite group, consequently*

$$B_\times(\mathcal{J}_a) = Q_\times(\mathcal{E}_a).$$

PROOF. The assertion is a direct consequence of Lemma 5, whence we obtain that

$$\bigcup_{1\le \nu < \frac{2\sqrt{a}}{\sqrt{3}}} \left(\nu Q_\times(\mathcal{J}_a) \cup \frac{1}{\nu} Q_\times(\mathcal{J}_a)\right) \quad \bigcup_{p|a} \left(p Q_\times(\mathcal{J}_a) \cup \frac{1}{p} Q_\times(\mathcal{J}_a)\right)$$

covers the group $Q_\times(\mathcal{E}_a)$. The proof is completed. ∎

REMARK. Similar theorem can be proved for $\psi_{a,A}(n) = An^2 + a$, if $a \in \mathbb{Z}$, $a \neq 0$, $A = 1, 2, 3, 4$.

Let

$$S_a := \frac{Q_\times(\mathcal{E}_a)}{Q_\times(\mathcal{J}_a)},$$

and for some $r \in Q_\times(\mathcal{E}_a)$, let $\overline{r} = r Q_\times(\mathcal{J}_a)$.

THEOREM 6. *For* $a \in [1,20]$, *we have*

$$S_a = \{\overline{1}\}, \qquad \textit{if } a = 1,2,4,7,8,16,18,$$
$$S_a = \{\overline{1},\overline{2}\}, \qquad \textit{if } a = 3,5,6,10,11,12,13,19,20,$$
$$S_a = \{\overline{1},\overline{3}\}, \qquad \textit{if } a = 9,14,17.$$

PROOF OF THEOREM 6. For $a \le 20$, $\dfrac{2\sqrt{a}}{\sqrt{3}} < 6$. Let L_a be the set of those numbers q from $2,3,5$ for which either $q|a$, or $\left(\dfrac{-a}{q}\right) = 1$. Then, by Lemma 5, we obtain immediately, that $Q_\times(\mathcal{J}_a \cup L_a) = Q_\times(\mathcal{E}_a)$. Thus, to prove the theorem, it is enough to decide, which elements of L_a do belong to $Q_\times(\mathcal{J}_a)$.

We proved that $Q_\times(\mathcal{E}_a) \ne Q_\times(\mathcal{J}_a)$ if $2^2 \| a+1$, (Lemma 8), or if $a \equiv 5 \pmod 8$ (Lemma 7), or if $a = bq$, where q is a prime, $\left(\dfrac{b}{q}\right) = -1$ (see Lemma 9).

We shall use Lemma 5.

For shortening we write $\mathcal{F} = Q_\times(\mathcal{J}_a)$, $\varphi(x) = x^2 + a$.

1. CASES $a = 1,4,9,16,25,36$. See Theorem 2.

2. CASE $a = 2$. $\varphi(1) = 3 \in \mathcal{F}$, $2 \in \mathcal{F}$, apply Lemma 5.

3. CASE $a = 3$. $3 \in \mathcal{F}$, $\varphi(1) = 4$, $\varphi(2) = 7$, $\varphi(5) = 28$, whence $2^2 \in \mathcal{F}$. Thus $Q_\times(\mathcal{E}_3) = \{\overline{1},\overline{2}\} \otimes Q_\times(\mathcal{J}_3)$.

4. CASE $a = 5$. See Theorem 4.

5. CASE $a = 6$. Since $\left(\dfrac{2}{3}\right) = -1$, from Lemma 9 we obtain that S_a contain at least two elements, furthermore that $q = 3 \notin \mathcal{F}$. Since $6 \in \mathcal{F}$, therefore $2 = 6/3$ is in the same residue class mod $Q(\mathcal{J}_6)$ as 3.

Since

$$\frac{\varphi(7)}{\varphi(2)} = \frac{11}{2}, \ \frac{\varphi(7)}{\varphi(2)} \cdot \varphi(4) = 11^2 \in \mathcal{F},$$

we obtain that $2^2 \in \mathcal{F}$.

Since $2\sqrt{\dfrac{6}{3}} = 2\sqrt{2} < 3$, therefore by Lemma 5 we obtain that each prime p, such that $\left(\dfrac{-a}{p}\right) = 1$, belongs to $Q_\times(\mathcal{I}_a) \cup 2Q_\times(\mathcal{I}_a)$.

Hence the assertion follows.

6. CASE $a = 7$. Since $\varphi(1) = 2^3$, $\varphi(3) = 11$, therefore $2 \in \mathcal{F}$. $3 \nmid \varphi(n)$, since $\left(\dfrac{-7}{3}\right) = \left(\dfrac{2}{3}\right) = -1$. We can apply Lemma 5, and deduce the

$$Q_\times(\mathcal{I}_a) \cup 2Q_\times(\mathcal{I}_a) = Q_\times(\mathcal{E}_a).$$

7. CASE $a = 8$. $\varphi(1) = 9$, $\varphi(2) = 12$, $\varphi(4) = 24$, thus $2 = \dfrac{\varphi(4)}{\varphi(2)} \in \mathcal{F}$, $3 = \dfrac{\varphi(4)}{2^3} \in \mathcal{F}$. We can apply Lemma 5.

8. CASE $a = 9$. See Theorem 2.

9. CASE $a = 10$. Since $\left(\dfrac{2}{5}\right) = -1$, therefore $5 \notin \mathcal{F}_{10}$. Thus $2 = \dfrac{10}{5} \notin \mathcal{F}$, $5 \equiv 2 \pmod{\mathcal{F}}$. Since $2\sqrt{\dfrac{a}{3}} < 4$, from Lemma 5 we obtain the assertion, stated.

10. CASE $a = 11$. We have $2^2 \| a + 1$, thus by Lemma 8, $2 \notin Q_\times(\mathcal{I}_{11})$. Since $\varphi(1) = 12$, $\varphi(5) = 36$, we have $3 \in Q_\times(\mathcal{I}_{11})$, $\dfrac{\varphi(2)}{3} = 5 \in Q_\times(\mathcal{I}_{11})$.

Apply Lemma 5.

11. CASE $a = 12$. Observe that, if $2^\alpha \| \varphi(n)$, then $\alpha =$ even, thus $2 \notin \mathcal{F}$. $\dfrac{\varphi(4)}{\varphi(3)} = \dfrac{28}{21} = \dfrac{4}{3} \in \mathcal{F}$, $\dfrac{4}{3} \cdot \dfrac{1}{a} = \dfrac{1}{3^2} \in \mathcal{F}$, $3^2 \in \mathcal{F}$, $\dfrac{\varphi(10)}{\varphi(4)} = \dfrac{112}{28} = 4 \in \mathcal{F}$, $\dfrac{a}{4} = 3 \in \mathcal{F}$, $\dfrac{\varphi(3)}{3} = 7 \in \mathcal{F}$.

Since $2\sqrt{\dfrac{a}{3}} = 2 \cdot 2 < 5$, therefore, by Lemma 5 we obtain the assertion.

12. CASE $a = 13$. Since $a \equiv 5 \pmod 8$, therefore $2 \notin \mathcal{F}$. Since $\dfrac{\varphi(8)}{\varphi(6)} = \dfrac{7 \cdot 11}{7^2} = \dfrac{11}{7} \in \mathcal{F}$, $7 \cdot 11 \in \mathcal{F}$, we obtain that $11^2 \in \mathcal{F}$, and that $7^2 \in \mathcal{F}$. Since $\varphi^2(1) = 2^2 \cdot 7^2$, therefore $2^2 \in \mathcal{F}$.

To finish the proof, use Lemma 5.

13. CASE $a = 14$. $14 \in \mathcal{F}$, $\varphi(1) = 15 \in \mathcal{F}$, $\varphi(4) = 30 \in \mathcal{F}$, $\dfrac{\varphi(4)}{\varphi(1)} = 2 \in \mathcal{F}$, $\dfrac{a}{2} = 7 \in \mathcal{F}$, $\dfrac{\varphi(2)}{2} = 3^2 \in \mathcal{F}$, $\varphi(6) = 2 \cdot 25$, $5^2 \in \mathcal{F}$, $\varphi(18) = 18^2 + 14 = 2 \cdot [2 \cdot 9^2 + 7] = 2 \cdot [169] \Rightarrow 13^2 \in \mathcal{F}$, $\varphi(20) = 414 = 2 \cdot 207 = 2 \cdot 9 \cdot 23$, $23 \in \mathcal{F}$.

Since $14, 15, 30 \in \mathcal{F}$, therefore $\dfrac{30}{15} = 2 \in \mathcal{F}$, $7 = \dfrac{14}{2} \in \mathcal{F}$. Furthermore $18 \in \mathcal{F}$, $18/2 = 3^2 \in \mathcal{F}$. By using Lemma 5, we obtain that each $p \in \mathcal{E}_{14}$ can be written as

$$p = 2^\alpha \cdot 3^\beta \rho,$$

where $\rho \in Q_\times(\mathcal{E}_a)$. Thus $S_{14} = \{\overline{1}\}$ if $3 \in Q_\times(\mathcal{E}_{14})$, and $S_{14} = \{\overline{1}, \overline{3}\}$, if $3 \notin Q_\times(\mathcal{E}_{14})$. We shall prove that the last assertion is true.

Assume that $3 \in Q_\times(\mathcal{E}_{14})$, i.e. that

$$3 = \prod_{i=1}^{h} (n_i^2 + 14)^{\varepsilon_i}, \qquad \varepsilon_i \in \{-1, 1\}. \tag{7.1}$$

It is clear that $7 | n^2 + 14$ if and only if $7|n$, and $(7m)^2 + 14 = 7\{7m^2 + 2\}$, $(7m^2 + 2, 7) = 1$.

We can rewrite (7.1) in the following form:

$$3 = \prod_{i=1}^{t} \left(n_i^2 + 14\right)^{\varepsilon_i} \cdot \prod_{j=1}^{s} 7^{\delta_j} \left(7m_i^2 + 2\right)^{\delta_j}$$

where $(n_i, 7) = 1$, $\varepsilon_i, \delta_j \in \{-1, 1\}$. Furthermore, $\delta_1 + \ldots + \delta_s = 0$, consequently

$$3 = \prod (n_i^2 + 14)^{\varepsilon_i} \prod (7m_i^2 + 2)^{\delta_j}.$$

Thus, by the Legendre symbol $\left(\frac{\cdot}{7}\right)$, we have

$$\left(\frac{3}{7}\right) = \prod \left(\frac{n_i^2}{7}\right)^{\varepsilon_i} \cdot \prod \left(\frac{2}{7}\right)^{\delta_j}.$$

Since $\sum \delta_j = 0$, $\left(\frac{n_i^2}{7}\right) = 1$, we obtain that $\left(\frac{3}{7}\right) = 1$, but this is not true.

The assertion is proved.

14. CASE 15. From Lemma 9, with $q = 5$, $b = 3$, $\left(\left(\frac{3}{5}\right) = -1\right)$ we obtain that $5 \notin \mathcal{F}$, consequently $3 \notin \mathcal{F}$. Since $\varphi(1) = 16$, $\varphi(3) = 24$, we obtain that $\frac{\varphi(3)}{\varphi(1)} = \frac{3}{2} \in \mathcal{F}$, $\varphi(7) = 64 = 2^6$, $\frac{2^6}{2^4} = 2^2 \in \mathcal{F}$, $\frac{15 \cdot \varphi(5)}{\varphi(3)} = 5^2 \in \mathcal{F}$, $\left(\frac{3}{2}\right)^2 \cdot 2^2 = 3^2 \in \mathcal{F}$. Thus $\overline{2} = \overline{3}$. By Lemma 5 we deduce that S_{15} has two elements.

The assertion is proved.

15. CASE 17. We obtain that $\varphi(1) = 18$, $\varphi(2) = 21$, $\varphi(5) = 42$, and so $\frac{\varphi(5)}{\varphi(2)} = 2 \in \mathcal{F}$, $3^2 \in \mathcal{F}$. We shall prove that $3 \notin Q_\times(\mathcal{J}_a)$. Assume indirectly that

$$3 = \prod_{i=1}^{h} \left(n_i^2 + 17\right)^{\varepsilon_i} \cdot \prod_{j=1}^{k} \left((17m_i)^2 + 17\right)^{\delta_j},$$

where $(n_i, 17) = 1$. Since

$$(17m_i)^2 + 17 = 17(17m_i^2 + 1),$$

we obtain that

$$3 = A \cdot B \cdot 17^\gamma, \tag{6.2}$$

where

$$\gamma = \sum_{j=1}^{k} \delta_j, \quad A = \prod \left(n_i^2 + 17\right)^{\varepsilon_i},$$

$$B = \prod_{j=1}^{k} \left(17m_i^2 + 1\right)^{\delta_j}.$$

Let us consider the Legendre-symbol $\left(\frac{\cdot}{17}\right)$. Since $(A, 17) = (B, 17) = (3, 17) = 1$, therefore $\gamma = 0$, furthermore

$$\left(\frac{3}{17}\right) = \left(\frac{A}{17}\right)\left(\frac{B}{17}\right).$$

Since

$$\left(\frac{A}{17}\right) = \prod\left(\frac{n_i^2}{17}\right) = 1, \quad \left(\frac{B}{17}\right) = \prod\left(\frac{1}{17}\right) = 1,$$

we deduce that

$$\left(\frac{3}{17}\right) = 1,$$

but this is not true.

Hence we obtain easily the assertion.

16. CASE 18. Since $18 \in \mathcal{F}$, $\varphi(2) = 22 \in \mathcal{F}$, $\varphi(3) = 27 \in \mathcal{F}$, we obtain that $\frac{\varphi(3)}{18} = \frac{3}{2} \in \mathcal{F}$. Furthermore $\frac{\varphi(6)}{a} = \frac{54}{18} = 3 \in \mathcal{F}$, and so $2 \in \mathcal{F}$. By using Lemma 5, we deduce that S_{18} is trivial.

17. CASE 19. From Lemma 8 we obtain that $2 \notin Q_\times(\mathcal{I}_{19})$. Furthermore, $\varphi(11) = 140$, $\varphi(1) = 20$, hence $7 \in \mathcal{F}$, $\frac{\varphi(3)}{7} = 2^2 \in \mathcal{F}$, $\frac{\varphi(4)}{7} = 5 \in \mathcal{F}$, $\frac{\varphi(3)}{18} = \frac{3}{2} \in \mathcal{F}$.

Thus $\overline{2} = \overline{3}$, and by Lemma 5, the assertion follows.

18. CASE 20. From Lemma 6 we obtain that $5 \in \mathcal{F}$, $2^2 \in \mathcal{F}$. Furthermore, $\varphi(4)/2^2 = 3^2 \in \mathcal{F}$. Hence, by Lemma 5 it is clear that S_{20} contains one or two elements. We shall prove that $2 \notin \mathcal{F}_{14}$.

Let us assume indirectly that $2 \in \mathcal{F}_{14}$. Then

$$2 = \prod \varphi(n_j)^{\varepsilon_j} = U_1 \cdot U_2 \cdot U_3$$

where

$$U_1 = \prod_{2^2 | n_j} \varphi^{\varepsilon_j}(n_j),$$

$$U_2 = \prod_{2 \| n_j} \varphi^{\varepsilon_j}(n_j),$$

$$U_3 = \prod_{2 \nmid n_j} \varphi^{\varepsilon_j}(n_j).$$

Since $\varphi(n) \equiv 1 \pmod 4$, if $(n, 2) = 1$, therefore $U_3 \equiv 1 \pmod 4$. Furthermore $\varphi(4m) = 4\left((2m)^2 + 5\right)$, $\dfrac{\varphi(4m)}{4} \equiv 1 \pmod 1$, and so $U_1 = 2^{2\sum_1} . V_1$, where $\sum_1 = \sum\limits_{2^2|n_j} \varepsilon_j$, $V_1 \equiv 1 \pmod 4$. Finally,

$$\varphi(2(2k+1)) = 2^2 \left\{(2k+1)^2 + 5\right\} = 2^2\{6 + 4k(k+1)\} = 2^3\{3 + 2k(k+1)\},$$

and so $\dfrac{\varphi(2(2k+1))}{2^3} \equiv -1 \pmod 4$.

Thus $U_2 = 2^{2\sum_2} \cdot V_2$, where $\sum_2 = \sum\limits_{2\|n_j} \varepsilon_j$, $V_2 \equiv (-1)^{\sum_2} \pmod 4$.

Consequently

$$2 = 2^{2\sum_1 + 3\sum_2} \cdot V_1 \cdot V_2 \cdot U_3,$$

$(V_1 V_2 U_3, 2) = 1$. Thus $2\sum_1 + 3\sum_2 = 1$, and so $\sum_2 =$ odd. Hence $1 = V_1 V_2 U_3$, $1 \pmod 4) \equiv (-1)^{\sum_2} \pmod 4$, which is a contradiction.

The assertion is proved. ∎

References

[1] I. KÁTAI, On sets characterizing number-theoretical functions, *Acta Arithmetica* **13** (1968), 315–320.

[2] D. WOLKE, Bemerkungen über Eindeutigkeitsmengen additiver Funktionen, *Elem. der Math.* **33** (1978), 14–16.

[3] P. D. T. A. ELLIOTT, *Arithmetic functions and integer products*, Springer V., Berlin, 1985.

[4] F. DRESS and B. VOLKMAN, Ensembles d'unicité pour les fonctions arithmetiques additives on multiplicatives, *C. R. Acad. Sci. Paris, Ser. A.* **287** (1978), 43–46.

[5] K.-H. INDLEKOFER, On sets characterizing additive and multiplicative arithmetical functions, *Illinois Journal of Math.* **25**, (1981), 251–257.

[6] K.-H. INDLEKOFER, On sets characterizing additive arithmetical functions, *Math. Z.* **146** (1976), 285–290.

[7] P. HOFFMANN, Note on a problem of Kátai, *Acta Math. Hung.* **45** (1985), 261–262.

[8] K.-H. INDLEKOFER, J. FEHÉR and L. STACHO, On sets of uniqueness for completely additive arithmetic functions, *Analysis* **16** (1996), 405–415.

[9] J. FEHÉR, K.-H. INDLEKOFER and N. M. TIMOFEEV, A set of uniqueness for completely additive arithmetic functions, *Annales Univ. Sci. Budapest, Sect. Comp.* **21** (2002), 57–67.

[10] J. M. DE KONINCK and I. KÁTAI, *On the multiplicative group generated by shifted binary quadratic form* (manuscript).

[11] K.-H. INDLEKOFER and N. M. TIMOFEEV, (unpublished)

J. Fehér
University of Pécs
Department of Mathematics
H-7624 Pécs
Ifjúság útja 6.
Hungary
feher@ttk.pte.hu

I. Kátai
Department of Computer Algebra,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
and
Research Group of Applied Number Theory
Hungarian Academy of Sciences
katai@compalg.inf.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 17–28

ON THE MULTIPLICATIVE GROUP GENERATED BY SHIFTED BINARY QUADRATIC FORMS

By

J. M. DE KONINCK* and I. KÁTAI

Dedicated to the memory of Nikolai M. Timofeev

(Received April 23, 2003)

1. Introduction

Let E be a set of positive integers. We say that E is a set of uniqueness modulo 1 if for each completely additive function $f : \mathbf{N} \to \mathbf{R}/\mathbf{Z}$ for which $f(e) \equiv 0 \pmod 1$ for every $e \in E$, we necessarily have that $f(n) \equiv 0 \pmod 1$ for each positive integer n. Here and in what follows, we let $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ stand for the set of positive integers, all integers, rational numbers and real numbers, respectively; also p always stands for a prime number. It is clear that the domain of a completely additive function f can be extended to the multiplicative group of positive rationals, simply by setting

$$f(m/n) = f(m) - f(n) \quad \text{for each } m, n \in \mathbf{N}.$$

Let $\mathbf{Q}^*$ be the multiplicative group of positive rationals, and for each positive integer h, let

$$Q_h^* := \left\{ \frac{m}{n} : m, n \in \mathbf{N},\ (mn, h) = 1 \right\}.$$

Let E^* be the multiplicative group generated by E. It was proved independently by several authors that E is a set of uniqueness mod 1 if and only if $E^* = \mathbf{Q}^*$; see for instance Indlekofer [5], Hoffman [3], Elliott [4] and Meyer [9]. It is not known whether the set of shifted primes is a set of uniqueness mod 1.

In Kátai [7], it was proved implicitly that the set of "primes + one" enlarged by a suitable finite set of primes is a set of uniqueness mod 1.

---

* Research supported in part by a grant from NSERC.

Mathematics Subject Classification (2000): 11E16, 11E25, 11H55, 11J71, 11K65

Elliott [2] proved that the set of primes up to 10^{387} together with the set of shifted primes forms a set of uniqueness mod 1.

Given a binary quadratic form $ax^2+bxy+cy^2$, let $-D=4ac-b^2$ stand for its discriminant. Now assume that D is equal to 4 or 8 or an odd prime. Let $\chi_D(n)=(\frac{-D}{n})$ be the Kronecker character and $\mathcal{B}(D)$ be the multiplicative semigroup generated by the union of the following four sets:

$$\{p : p|D\}, \qquad \{r^2 : r=1,2,3,\ldots\}, \qquad \{p : \chi_{-D}(p)=1\}, \qquad \{0\}.$$

From here on, when the context is clear, we shall write χ instead of χ_D. Now let

$$w(n) := \sum_{d|n}\chi(d) = \prod_{p^\alpha\|n}\left(1+\chi(p)+\ldots+\chi(p^\alpha)\right). \tag{2.1}$$

It is clear that an integer n coprime to D belongs to $\mathcal{B}(D)$ if and only if $w(n)>0$. Furthermore, if $(n,D)=1$, then it is well known that the number of representations of n by classes of binary quadratic forms with discriminant $-D$ is $\alpha w(n)$, where

$$\alpha = \begin{cases} 2 & \text{if } D>4,\\ 4 & \text{if } D=4,\\ 6 & \text{if } D=3 \end{cases}$$

(see Landau [8]). Assume that A is a positive integer and set

$$E(D,A) := \{n+A : n\in\mathcal{B}(D)\}.$$

Let furthermore $\mathcal{H}(D,A)$ be the multiplicative group generated by $E(D,A)$.

In this paper, we study the set $\mathcal{H}(D,A)$ in the cases where D is 4, 8 or an arbitrary prime number larger than 3.

REMARKS.

(a) Fehér, Indlekofer and Timofeev [6] investigated the case $D=4$ and proved that $\mathcal{H}(4,A)=\mathbf{Q}^*$, if A is the sum of two squares.

(b) Indlekofer and Timofeev proved a more precise result for the set $\{n+A \mid n\in\mathcal{B}(4)\}$ [10], namely that for $a,\ b\in\mathbf{N}$ $(a,b)=1$ $(ab,2A)=$ $=1$, there are infinitely many $n,\ m\in\mathcal{B}(4)$, such that $a(n+A)=b(m+A)$, $(a,n+A)=1$.

(c) If the class number $h(-D)=1$, then $D=4,8$ or an odd prime, and $\mathcal{B}(D)$ can be interpreted as the set of those integers which can be written as the values of one binary quadratic form of discriminant $-D$.

2. Main results

THEOREM 1. *Let $D > 3$ be an arbitrary prime and let A be any given positive integer. Then*

$$\mathcal{H}(D, A) = \begin{cases} \mathbf{Q}_D^* & \text{if } \chi_D(A) = -1, \\ \mathbf{Q}^* & \text{otherwise.} \end{cases}$$

THEOREM 2. *Let $D = 4$ and let A be an arbitrary positive integer. Then $\mathcal{H}(4, A) = \mathbf{Q}^*$.*

THEOREM 3. *Let $D = 8$ and let A be an arbitrary positive integer. Then $\mathcal{H}(8, A) = \mathbf{Q}^*$.*

3. Preliminary lemmas

LEMMA 0. *Let χ be the Kronecker character mod $-D$, where $D > 0$. Let $U > 0$ and $V \neq 0$ be two integers for which there is an arithmetic progression $\ell \pmod{D}$ such that $\chi(\ell) = 1$ and such that $t := U\ell + V$ satisfies $\chi(t) = 1$. Moreover, let*

$$a(x) := \sum_{\substack{x < p \leq 2x \\ p \equiv \ell \pmod{D}}} w(Up + V),$$

where w is defined by (2.1). Then $a(x)$ is positive if x is sufficiently large.

We shall not prove this lemma. Indeed, the result can easily be obtained by using the Bombieri–Vinogradov mean value theorem in the form

$$\sum_{k \leq \sqrt{x}/(\log x)^{B+25}} \max_{\ell} \max_{n \leq x} \left| \pi(u, k, \ell) - \frac{\mathrm{li}(x)}{\phi(k)} \right| \ll \frac{x}{\log^B x}$$

(see Elliott [1], Chapter 7), where $\mathrm{li}(x)$ stands for the logarithmic integral, and the "enveloping sieve" given by Hooley (see [4], Chapter 5), which he used to obtain an asymptotic estimate for the number of solutions of the equation $n = p + x^2 + y^2$. A more detailed argument is given in the proof of Lemma 1.

In the following lemmas, we assume that D is an odd prime and $(A, D) = 1$.

LEMMA 1. *Let* $k \equiv 1 \pmod D$ *and* $(k, A) = 1$. *Then* $k \in \mathcal{H}(D, A)$.

PROOF. In order to prove that $k \in \mathcal{H}(D, A)$, it is sufficient to find $n_1, n_2 \in \mathcal{B}(D)$ such that $n_1 + A = k(n_2 + A)$. Let p run over the set of primes $p \equiv 1 \pmod D$ (so that $p \in \mathcal{B}(D)$) and consider the sum

$$a(x) := \sum_{x<p\le 2x} w(kp + (k-1)A).$$

It is enough to prove that $a(x)$ is positive for some x.

To do so, we let $\ell(p) := kp + (k-1)A$ and observe that $\ell(p) \equiv 1 \pmod D$, so that $\chi(\frac{\ell(p)}{d}) = \chi(d)$. Consequently, using the definition of w given in (2.1), we have

$$w(\ell(p)) = 2 \sum_{\substack{d \mid \ell(p) \\ d<\sqrt{\ell(p)}}} \chi(d) + E_p,$$

where $E_p = 0$ except when $\ell(p)$ is a square, in which case $E_p = \chi\left(\sqrt{\ell(p)}\right)$, that is $|E_p| \le 1$.

Thus, given a large number B,

$$\begin{aligned} a(x) = &\sum_{d\le \sqrt{x}/\log^B x} 2\chi(d)\cdot \#\{p \in [x, 2x] : \ell(p) \equiv 0 \pmod d\} \\ &+ \sum_{\sqrt{x}/\log^B x<d\le\sqrt{2kx+(k-1)A}} 2\chi(d)\cdot \#\{p \in [x, 2x] : \ell(p) \equiv 0 \pmod d,\ d^2 < \ell(p)\} \\ &+ O(\sqrt{x}) \\ &= \Sigma_1 + \Sigma_2 + O(\sqrt{x}), \end{aligned}$$

say. Using the Bombieri–Vinogradov mean value theorem (stated above), one can obtain that

$$\Sigma_1 = 2\,(\mathrm{li}(2x) - \mathrm{li}(x)) \sum_{d\le\sqrt{x}/\log^B x} \frac{\chi(d)}{\phi(dD)} + O\left(\frac{x}{\log^{B_1} x}\right),$$

where ϕ stands for the Euler function and where B_1 can be taken arbitrarily large provided B is large enough.

The crucial step is the evaluation of Σ_2. This can be done by using the *enveloping sieve* of Hooley. We shall not go into details, but one can easily obtain from this method that, as $x \to \infty$,

$$a(x) = C(D)\frac{2x}{\log x} + o\left(\frac{x}{\log x}\right),$$

where $C(D) = \sum_{d=1}^{\infty} \frac{\chi(d)}{\phi(dD)}$, which proves Lemma 1. ∎

LEMMA 2. *Let* $k \equiv \ell \pmod{D}$ *and* $(k\ell, AD) = 1$. *Then* $k/\ell \in \mathcal{H}(D, A)$.

PROOF. Since both $k\ell^{\phi(D)-1}$ and $\ell^{\phi(D)}$ are $\equiv 1 \pmod{D}$ and are coprime to A, then they both belong to $\mathcal{H}(D, A)$, from which it follows that their ratio k/ℓ also belongs to $\mathcal{H}(D, A)$. ∎

LEMMA 3. *Let* $\mathbf{Z}_D^*$ *be the set of reduced residue classes mod* D *and let* $\mathcal{I}$ *be its subgroup generated by*

$$\{\nu + A : \nu = 0 \text{ or } \nu = \text{ quadratic residue mod } D\} \setminus \{0\}. \tag{3.1}$$

Then $\mathcal{I} = \mathbf{Z}_D^*$.

PROOF. Assume that $\mathcal{I}$ is a proper subgroup of $\mathbf{Z}_D^*$. Then $\#\mathcal{I} < D - 1$, so that $\#\mathcal{I} \le (D-1)/2$. On the other hand, since the set of its generating elements contains $(D-1)/2$ members, then $\#\mathcal{I}$ must be equal to $(D-1)/2$, so that $\mathcal{I}$ must be the subgroup of the quadratic residues mod D. This means that $\nu + A$ is a quadratic residue if ν is equal to zero or to a quadratic residue, except when $\nu = -A$. (Observe that, in the case $\chi(-A) = -1$, $\mathcal{I}$ always has at least $(D+1)/2$ elements, so that $\#\mathcal{I} = D - 1$, in which case $\mathcal{I} = \mathbf{Z}_D^*$.) Thus

$$\sum_{m=0}^{D-1} (\chi(m) + 1)(\chi(m + A) + 1) \ge 2 + 4 \cdot \frac{D-3}{2}. \tag{3.2}$$

But, since

$$\sum_{m=0}^{D-1} \chi(m) = \sum_{m=0}^{D-1} \chi(m + A) = 0 \quad \text{and} \quad \sum_{m=0}^{D-1} \chi(m)\chi(m + A) = -1,$$

it follows that the left hand side of (3.2) is $D - 1$ and therefore that $D - 1 \ge$ $\ge 2 + 4 \cdot \frac{D-3}{2}$, which is impossible if $D > 3$.

It remains to consider the case $D = 3$. If $A \equiv 1 \pmod 3$, then the set $\{0+1 \pmod 3, 1+1 \pmod 3\}$ generates $\mathbf{Z}_3^*$. If $A \equiv -1 \pmod 3$, then $(-1) \pmod 3 \in \mathcal{J}$ and $(-1)^2 \pmod 3 \in \mathcal{J}$, so that $\mathcal{J} = \mathbf{Z}_3^*$, thus completing the proof of Lemma 3. ∎

LEMMA 4. *Let $\chi(-A) = -1$. Then $\mathcal{H}(D, A) \subseteq \mathbf{Q}_D^*$.*

PROOF. It is enough to show that $(n + A, D) = 1$ for every $n \in \mathcal{B}(D)$. Indeed, if $n + A \equiv 0 \pmod D$, then $\chi(n) \equiv \chi(-A) = 1$ and consequently $(n, D) = 1$. But $n \in \mathcal{B}(D)$ and $(n, D) = 1$ imply that $\chi(n) = 1$, thus proving Lemma 4. ∎

LEMMA 5. *Let S_A be the multiplicative group generated by $E_1 \cup E_2$, where*

$$E_1 = \{p + A : \chi(p) = 1,\ p \not\equiv -A \pmod D\},$$
$$E_2 = \{D^r + A : r = 1, 2, 3, \dots\}.$$

Then, for every $v \in \mathbf{Z}_D^*$, S_A contains infinitely many integers congruent to $v \pmod D$, all of which are coprime to A. Moreover, $S_A \subseteq \mathcal{H}(D, A)$.

PROOF. These results are direct consequences of Lemma 3. ∎

4. Proof of Theorem 1

Assume first that $(A, D) = 1$. Then it follows from Lemmas 2 and 5 that

$$\mathbf{Q}_{AD}^* \subseteq \mathcal{H}(A, D).$$

Let $A = \pi_1^{\alpha_1} \pi_2^{\alpha_2} \dots \pi_r^{\alpha_r}$. We shall prove that $\pi_j \in \mathcal{H}(A, D)$ for $j = 1, 2, \dots, r$, which will imply that

$$\mathbf{Q}_D^* \subseteq \mathcal{H}(A, D). \tag{4.1}$$

So let π_1 be one of the prime divisors of A and write $A = \pi_1^{\alpha_1} A_2$.

Assume first that $\alpha_1 = 1$. Then for $m \in \mathcal{B}(D)$, we have

$$\mathcal{H}(A, D) \ni \pi_1^2 Dm + A = \pi_1 (\pi_1 Dm + A_2).$$

Since $(\pi_1 Dm + A_2, AD) = 1$, it follows that $\pi_1 Dm + A_2 \in \mathcal{H}(A, D)$, and so $\pi_1 \in \mathcal{H}(A, D)$.

For $\alpha_1 > 1$, we consider separately the cases α_1 odd and α_1 even.

First assume that $\alpha_1 = 2\beta + 1$, with $\beta \geq 1$. Then we have

$$\pi_1^{2\beta+2} Dm + \pi_1^{2\beta+1} A_2 = \pi_1^{2\beta+1} (\pi_1 Dm + A_2) \in \mathcal{H}(A, D).$$

Since $(\pi_1 Dm + A_2, AD) = 1$, we obtain that $\pi_1 Dm + A_2 \in \mathcal{H}(A, D)$ and consequently that $\pi_1^{2\beta+1} \in \mathcal{H}(A, D)$. Furthermore, if $m \in \mathcal{B}(D)$, then $\pi_1^2 Dm + A \in \mathcal{H}(A, D)$ and $\pi_1^2 Dm + A = \pi_1^2(Dm + \pi_1^{2\beta-1} A_2)$, whence $\pi_1^2 \in \mathcal{H}(A, D)$ follows by observing that $(Dm + \pi_1^{2\beta-1} A_2, AD) = 1$. Thus

$$\pi_1 = \frac{\pi_1^{2\beta+1}}{(\pi_1^2)^\beta} \in \mathcal{H}(A, D).$$

Let us now consider the case $\alpha = 2\beta$ with $\beta \geq 1$. Starting with $m \in$ $\in \mathcal{B}(D)$, since

$$\mathcal{H}(A, D) \ni \pi_1^{2\beta+2} Dm + A = \pi_1^{2\beta} \left(D\pi_1^2 m + A_2 \right),$$

$(D\pi_1^2 m + A_2, AD) = 1$, it follows that $D\pi_1^2 m + A_2 \in \mathcal{B}(D)$, and therefore that $\pi_1^{2\beta} \in \mathcal{B}(D)$.

We shall now prove that $\pi_1^2 \in \mathcal{H}(A, D)$. Since we already proved this in the case $\beta = 1$, we may assume that $\beta \geq 2$ and consider the integer $\pi_1^2 D +$ $+ A = \pi_1^2 \left(D + \pi_1^{2(\beta-1)} A_2 \right)$. Since $\pi_1^2 D + A \in \mathcal{H}(A, D)$, $D + \pi_1^{2(\beta-1)} A_2 \in$ $\in \mathcal{H}(A, D)$, we obtain that $\pi_1^2 \in \mathcal{H}(A, D)$, as claimed.

Finally, we observe that there is some $m \in \mathcal{B}(D)$ such that $\pi_1 \| mD +$ $+ A_2$. This is true if $Dm + A_2 \equiv \pi_1 \pmod{\pi_1^2}$, which defines an arithmetic progression $m \equiv s \pmod{\pi_1^2}$, where $s = (\pi_1 - A_2)D^{-1} \pmod{\pi_1^2}$, $(s, \pi_1) =$ $= 1$. If m is a prime p satisfying $p \equiv s \pmod{\pi_1^2}$ and $p \equiv 1 \pmod{D}$, then it is a suitable choice for $m \in \mathcal{B}(D)$, $\pi_1 \| Dm + A_2$.

Hence $Dm + A_2 = \pi_1 \eta$ with $(\eta, DA) = 1$ and $\eta \in \mathcal{H}(A, D)$; furthermore, $\pi_1^{2\beta} Dm + A = \pi_1^{2\beta}(Dm + A_2)$. Thus $\pi_1 \in \mathcal{H}(A, D)$ and since π_1 was an arbitrary prime divisor of A, our claim (4.1) is established.

Let us now investigate whether D belongs to $\mathcal{H}(A, D)$ or not. Since we already proved that it does not if $\chi(-A) = -1$, we may assume that $\chi(-A) = 1$. Then $p \equiv -A \pmod{D}$ implies that $p + A \in \mathcal{H}(A, D)$. There are infinitely many primes p such that $D \| p + A$, that is $\frac{p+A}{D} = \eta_p$ with $(\eta_p, D) = 1$

and $\eta_p \in \mathcal{H}(A, D)$, and consequently $D \in \mathcal{H}(A, D)$. Thus the theorem is proved in the case $(A, D) = 1$. Hence we shall now assume that $A = D^r B$ with $(B, D) = 1$ and $r \geq 1$. We shall try to find integers $n_1, n_2 \in \mathcal{B}(D)$ such that $n_1 + A = D(n_2 + A)$, that is $n_1 - Dn_2 = (D-1)A$. We shall find these by looking for m_1, m_2's such that $n_1 = D^r m_1$, $n_2 = D^{r-1} m_2$, which leads to the equation

$$m_1 - m_2 = (D-1)B. \tag{4.2}$$

Let ν run over zero and the quadratic residues mod D, that is over $\frac{D+1}{2}$ integers, and let $(H, D) = 1$. Then the set $\{\nu + H\}$ contains either a quadratic residue or zero. This is true in particular if we choose $H = (D-1)B$. So let ν, μ be such a couple of residues for which

$$\nu - \mu = (D-1)B, \qquad \chi(\nu) \neq -1, \qquad \chi(\mu) \neq -1.$$

If $\mu \not\equiv 0 \pmod{D}$, consider the sum

$$\sum_{\substack{x < p \leq 2x \\ p \equiv \mu \pmod{D}}} w(p + (D-1)B). \tag{4.3}$$

If $\mu \equiv 0 \pmod{D}$, then consider the sum

$$\sum_{\substack{x < p \leq 2x \\ p \equiv 1 \pmod{D}}} w(Dp + (D-1)B). \tag{4.4}$$

By using the Bombieri–Vinogradov mean value theorem and the evaluating sieve of Hooley mentioned in Lemma 0, one can deduce that both expressions (4.3) and (4.4) are positive provided x is large enough, in which case there exists at least one pair of integers $n_1, n_2 \in \mathcal{B}(D)$ for which

$$D = \frac{n_1 + A}{n_2 + A}.$$

The proof of Theorem 1 is thus complete. ∎

5. Proof of Theorem 2

Assume first that A is odd. We shall prove that

$$k = \frac{n_1 + A}{n_2 + A}, \qquad n_1, n_2 \in \mathcal{B}(4) \tag{5.1}$$

can be solved if $k \equiv 1 \pmod 4$, $(k, A) = 1$. Let n_2 run over the primes $p \equiv 1 \pmod 4$ and $n_1 = kp + (k-1)A$. By using the method of §4, one can prove that

$$\sum_{\substack{p<x \\ p\equiv 1 \pmod 4}} w(kp + (k-1)A) > 0$$

provided x is large enough, in which case (5.1) has a solution.

Hence we can deduce that for $k \equiv \ell \equiv 3 \pmod 4$, $(k\ell, A) = 1$, we have

$$k/\ell \in \mathcal{H}(4, A), \tag{5.2}$$

simply by repeating the argument used in the proof of Lemma 2.

Since $A+4, A+2 \in \mathcal{H}(4, A)$, there exists at least one $\nu \in \mathcal{H}(4, A)$ for which $\nu \equiv 3 \pmod 4$ and $(\nu, A) = 1$. Hence we obtain as earlier that

$$\mathbf{Q}_{4A}^{*} \subseteq \mathcal{H}(4, A).$$

Let $A = \pi_1^{\alpha_1} A_2$, $(A_2, \pi_1) = 1$, π_1 prime. We shall prove that $\pi_1 \in \mathcal{H}(4, A)$. Since π_1 is an arbitrary prime divisor of A, it will be true for each prime divisor of A, which implies that

$$\mathbf{Q}_4^{*} \subseteq \mathcal{H}(4, A). \tag{5.3}$$

Assume first that $\alpha_1 = 1$. Then $4\pi_1^2 + A_2\pi_1 = \pi_1(4\pi_1 + A_2)$ with $(4\pi_1 + A_2, 4A) = 1$, whence $\pi_1 \in \mathcal{H}(4, A)$.

Now consider the case $\alpha_1 = 2\beta + 1$, $\beta \geq 1$. By setting $4\pi_1^2 + A_2\pi_1^{2\beta+1} = \pi_1^2(4 + A_2\pi_1^{2\beta-1})$, we obtain that $\pi_1^2 \in \mathcal{H}(4, A)$. Then by considering $4\pi_1^{2\beta+2} + \pi_1^{2\beta+1}A_2 = \pi_1^{2\beta+1}(4\pi_1 + A_2)$ and observing that $4\pi_1 + A_2 \in \mathcal{H}(4, A)$, it follows that $\pi_1^{2\beta+1} \in \mathcal{H}(4, A)$, and hence that $\pi_1 \in \mathcal{H}(4, A)$.

Finally, let $\alpha = 2\beta$, $\beta \geq 1$. Similarly, by choosing the numbers $4\pi_1^{2\beta+2} + A$ and $4\pi_1^2 + A$, we first deduce that $\pi_1^2 \in \mathcal{H}(4, A)$.

Arguing as in the proof of Theorem 1, we first prove that there is at least one (actually infinitely many) $m \in \mathcal{B}(4)$ such that $Dm + A_2 \equiv \pi_1 \pmod{\pi_1^2}$. If such an integer m exists, then the integer $\eta_m = \frac{Dm+A_2}{\pi_1}$ is coprime to AD. Consequently $\eta_m \in \mathcal{H}(4, A)$ and furthermore $\pi_1^{2\beta+1}\eta_m = Dm\pi_1^{2\beta} + A \in \mathcal{H}(4, A)$, whence $\pi_1^{2\beta+1} \in \mathcal{H}(4, A)$, and so $\pi_1 \in \mathcal{H}(4, A)$.

It remains to prove the existence of such an integer m. To do so, it is enough to observe that there is at least one (actually infinitely many) prime $p \equiv 1 \pmod 4$ such that $4p + A_2 \equiv \pi_1 \pmod{\pi_1^2}$. Since this clearly holds, we have thus established (5.3).

We shall now prove that $2 \in \mathcal{H}(4, A)$.

If $A \equiv 1 \pmod 4$, then $2\|1 + A$ and $1 + A \in \mathcal{H}(4, A)$ imply that $2 \in \mathcal{H}(4, A)$.

If $A \equiv 3 \pmod 4$, then $A = -1+2^\gamma B$, with B odd and $\gamma \geq 2$. For every $\varepsilon > \gamma$, the number of primes $p < x$ for which $2^\varepsilon \| p+A$ is $(1+o(1))\mathrm{li}(x)/2^{\varepsilon-1}$, which means that there exists a prime p_ε and an odd integer $\eta_\varepsilon \in \mathcal{H}(4, A)$ such that $p_\varepsilon + A = 2^\varepsilon \eta_\varepsilon$. It is obvious that $p_\varepsilon \equiv 1 \pmod 4$ and thus that $p_\varepsilon + A \in \mathcal{H}(4, A)$. Hence

$$2 = \frac{2^{\varepsilon+1}}{2^\varepsilon} = \frac{p_{\varepsilon+1} + 1}{\eta_{\varepsilon+1}} \cdot \frac{\eta_\varepsilon}{p_\varepsilon + 1} \in \mathcal{H}(4, A).$$

We have thus proved that $\mathcal{H}(4, A) = \mathbf{Q}^*$ if $(A, 2) = 1$.

Assume now that $A = 2^\gamma B$ with B odd and $\gamma \geq 1$. We already proved that $\mathcal{H}(4, B) = \mathbf{Q}^*$, that is that each rational number m/n has a representation

$$\frac{m}{n} = \prod_{j=1}^{r} (n_j + B)^{\varepsilon_j},$$

where $\varepsilon_j \in \{-1, 1\}$ and $n_j \in \mathcal{B}(4)$, and so

$$\frac{m}{n} = 2^{\gamma(\varepsilon_1+\ldots+\varepsilon_r)} \prod_{j=1}^{r} (2^\gamma n_j + A)^{\varepsilon_j}.$$

To complete the proof of Theorem 2, it is enough to show that $2 \in \mathcal{H}(4, A)$. But this is true if

$$n_1 + A = 2(n_2 + A), \qquad n_1, n_2 \in \mathcal{B}(4)$$

can be solved. By writing $n_1 = 2^\gamma m_1$, $n_2 = 2^\gamma m_2$, it follows that the existence of $m_1, m_2 \in \mathcal{B}(4)$, with $m_1 - 2m_2 = B$, would be enough.

Now if $B \equiv 1 \pmod 4$, then let m_2 run over the set $\{2p : p \equiv 1 \pmod 4\}$ and consider the sum

$$\sum_{\substack{p \leq x \\ p \equiv 1 \pmod 4}} w(4p + B),$$

which is surely positive if x is large enough.

On the other hand, if $B \equiv -1 \pmod 4$, then let m_1 run over the set $\{2p : p \equiv 1 \pmod 4\}$ and consider the slightly different sum

$$\sum_{\substack{p \le x \\ p \equiv 1 \pmod 4}} w(2p + B),$$

which again is surely positive if x is large enough.

This completes the proof of Theorem 2. ∎

6. Proof of Theorem 3

Since the proof is very similar to that of Theorems 1 and 2, we shall only give a sketch of it.

Observe that now $D = 8$ and

$$\chi(1) = \chi(3) = 1, \quad \chi(5) = \chi(7) = -1.$$

Assume first that A is odd. Arguing as earlier, we can deduce that

$$\mathbf{Q}_{2A}^* \subseteq \mathcal{H}(8, A).$$

Repeating the argument used before, one can also prove that $\pi \in \mathcal{H}(8, A)$ if π is a prime divisor of A. Consequently,

$$\mathbf{Q}_2^* \subseteq \mathcal{H}(8, A).$$

Since $A + 1, A + 3 \in \mathcal{H}(8, A)$ and since either $2 \| A + 1$ or $2 \| A + 3$, we obtain that $2 \in \mathcal{H}(8, A)$, and so

$$\mathbf{Q}_4^* \subseteq \mathcal{H}(8, A).$$

The theorem is thus proved for A odd. So let $A = 2^\gamma B$ with B odd and $\gamma \ge 1$. As earlier, we can deduce that each rational number m/n can be written as

$$\frac{m}{n} = 2^{\Gamma(m,n)} \alpha(m, n),$$

where $\Gamma(m, n)$ is a positive integer depending on m and n, and $\alpha(m, n) \in$ $\in \mathcal{H}(8, A)$.

Thus it remains to prove that $2 \in \mathcal{H}(8, A)$. For this we try to solve the equation $n_1 + A = 2(n_2 + A)$, that is $n_1 - 2n_2 = A$. So let $n_1 = 2^\gamma m_1$, $n_2 = 2^\gamma m_2$, that is $m_1 - 2m_2 = B$. Let us now choose m_1 as follows

$$m_1 = \begin{cases} 2p + B & \text{with } p \equiv 1 \pmod 8 \text{ if } B \equiv 1 \pmod 8, \\ 2p + B & \text{with } p \equiv 3 \pmod 8 \text{ if } B \equiv 5 \pmod 8, \\ 8p + B & \text{with } p \equiv 1 \pmod 8 \text{ if } B \equiv 3 \pmod 8, \\ 2p + B & \text{with } p \equiv 1 \pmod 8 \text{ if } B \equiv 7 \pmod 8. \end{cases}$$

Since each of the above choices has at least one solution $m_1 \in \mathscr{B}(8)$, this completes the proof of Theorem 3. ∎

References

[1] P. D. T. A. ELLIOTT, *Arithmetic Functions and Integer Products*, Springer-Verlag, 1985.

[2] P. D. T. A. ELLIOTT, On representing integers as products of the $p+1$, *Monatsh. Math.* **97(2)** (1984), 85–97.

[3] P. HOFFMANN, Note on a problem of Kátai, *Acta Math. Hung.* **45** (1985), 261–262.

[4] C. HOOLEY, *Application of Sieve Methods to the Theory of Nunbers*, Cambridge University Press, 1976.

[5] K.-H. INDLEKOFER, On sets characterizing additive and multiplicative functions, *Ill. J. Math.* **25** (1981), 251–257.

[6] J. FEHÉR, K.-H. INDLEKOFER and N. M. TIMOFEEV, A set of uniqueness for completely additive arithmetic functions, *Annales Univ. Sci. Budapest, Sect. Comp.* **21** (2002), 57–67.

[7] I. KÁTAI, On sets characterizing number theoretical functions II. The set of "primes plus one" is a set of uniqueness; *Acta Arith.* **16** (1969/1970), 1–4.

[8] E. LANDAU, *Vorlesungen über Zahlentheorie, I–III.* S. Hirzel, Leipzig. English translation of Vol. I: Chelsea, New York, 1958.

[9] F. MEYER, Ensemble d'unicité pour les fonctions additives. Étude analogue dans le cas des fonctions multiplicatives; *Journées de Théorie Analytique et Élémentaire des nombres*, Orsay, 2 et 3 juin, 1980. *Publications Mathématiques d'Orsay*, 50–66.

[10] K.-H. INDLEKOFER and N. M. TIMOFEEV, Shifted B-numbers as a set of uniqueness for additive and multiplicative functions, *Acta Arithmetica* (accepted).

Jean-Marie De Koninck
Département de Mathématiques
Université Laval
Québec G1K 7P4
Canada
jmdk@mat.ulaval.ca

Imre Kátai
Department of Computer Algebra,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
katai@compalg.inf.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 29–33

ON A NEW PROOF OF A THEOREM OF INDLEKOFER AND TIMOFEEV

By

I. KÁTAI* and M. V. SUBBARAO†

(Received April 23, 2003)

1. Introduction

According to a theorem of H. Maier and G. Tenenbaum [1] (see [2], also) for all but $o(x)$ of integers $n \le x$ always there exist divisors $d_1 < d_2$ such that $d_2 < 2d_1$. See [2] as well.

Indlekofer and Timofeev [4] proved the same for the set of shifted primes. Namely they proved something more:

The number of primes $p \le x$, such that $p-1$ does not have a couple of divisors d_1, d_2 such that $d_1 < d_2 < 2d_1$, is less than

$$c\pi(x)\frac{(\log\log\log x)^{4\beta}}{(\log\log x)^{\beta}},$$

where $\beta = 1 - \dfrac{1+\log\log 3}{\log 3}$.

Our purpose here is to give a simple proof of the analogon of the theorem of Maier and Tenenbaum. In the proof we shall use only some simple sieve results and the Siegel–Walfisz theorem.

THEOREM (INDLEKOFER, TIMOFEEV). *With the exception of at most $o(lix)$ primes $p \le x$ there exist divisors $d_1 < d_2 (< 2d_1)$ of $p-1$.*

---

* The paper was written while the first named author visited Edmonton as a research professor invited by M. V. Subbarao funded by his NSERC grant. Financially supported also by the research group Applied Number Theory of the Hungarian Academy of Sciences.

† Supported by the NSERC grant of the author.

2. Proof

Let p, q with suffixes and without suffixes denote prime numbers, let $\mathcal{P}$ be the whole set of primes. Let $P(n)$ be the largest prime divisor of n.

Assume that $t = t(x) \to \infty$ very slowly, $t = O(\log\log\log x)$, say.

Let $\mathcal{B}_t = \{n,\ P(n) \le t\}$.

For some integer n let

$$M(n) := \prod_{\substack{p^\alpha \| n \\ p \le t}} p^\alpha, \quad E(n) = \prod_{\substack{p^\gamma \| n \\ p > t}} p^\gamma. \tag{2.1}$$

We know that

$$\pi(x, k, 1) \le c(\varepsilon) \frac{\operatorname{li} x}{\varphi(k)} \quad \text{for} \quad k \le x^{1-\varepsilon}, \tag{2.2}$$

and

$$\pi(x, k, 1) \le \frac{x}{k} \quad \text{if} \quad x^{1-\varepsilon} \le k \le x. \tag{2.3}$$

According to Lemma 5.2 in K. Prachar [3],

$$\begin{aligned} &\#\{n < x \mid P(n) > y\} < \\ &< x \exp\left(-\frac{\log_3 y}{\log y}\log x + \log_2 y + O\left(\frac{\log_2 y}{\log_3 y}\right)\right). \end{aligned} \tag{2.4}$$

From (2.4) we can get easily that

$$\limsup_{x\to\infty} \frac{1}{\pi(x)} \#\left\{p \le x \mid M(p+1) > t^\alpha\right\} < c(\alpha), \tag{2.5}$$

where $c(\alpha) \to 0$ if $\alpha \to \infty$.

Let α be fixed, $m < t^\alpha,\ m \in \mathcal{B}_t$.

Let

$$\prod\nolimits_m (x) := \#\{p \le x \mid M(p+1) = m\}. \tag{2.6}$$

Let

$$Q = \prod_{p<t} p.$$

We have

$$\prod\nolimits_m(x) = \#\left\{p \le x,\ m|p-1,\ \left(\frac{p-1}{m}, Q\right) = 1\right\},$$

and so

$$\begin{aligned}\prod\nolimits_m(x) &= \sum_{\delta|Q} \mu(\delta)\pi(x, \delta m; 1)\\ &= f(m)\mathrm{li}\, x + O\left(\frac{(\mathrm{li}\, x)}{\varphi(m)}(\log x)^{-A}\right),\end{aligned}$$

where

$$f(m) = \sum_{\delta|Q} \frac{\mu(\delta)}{\varphi(\delta m)} = \frac{1}{m}\prod_{\substack{\pi<t\\ \pi\nmid m}}\left(1 - \frac{1}{\pi-1}\right), \tag{2.7}$$

where π runs over $\mathcal{P}$.

Note that $2|m$ always holds. We can rewrite (2.7) as

$$f(m) = a(m)(1 + o_x(1))\frac{c^*}{\log t},$$

where

$$a(m) = \frac{1}{m}\prod_{\substack{\pi|m\\ \pi>2}} \frac{\pi-1}{\pi-2},$$

and c^* is defined from

$$\frac{c^*}{(\log t)}(1 + o_x(1)) = \prod_{2<p<t}\frac{\pi-2}{\pi-1}.$$

Let E_0 be the set of those integers n, for which no divisors d_1, d_2 exist with $d_1 < d_2 < 2d_1$. According to the Meier–Tenenbaum theorem,

$$\#(E_0 \cap [1, x]) < \rho(x)x, \tag{2.8}$$

where $\rho(x) \downarrow 0$ as $x \to \infty$.

Let

$$T(x) := \#\{p \le x \mid p+1 \in E_0\}.$$

Then

$$T(x) \le \sum_{m\in E_0}\prod\nolimits_m(x) = \sum\nolimits_1 + \sum\nolimits_2,$$

where in $\sum_1$, $m \le t^\alpha$, and in $\sum_2 : m > t^\alpha$.

As we noted, $\sum_2 \leq 2c(\alpha)\mathrm{li}\, x$, if x is large.

We have

$$\sum\nolimits_1 \leq \frac{2C^*}{(\log t)} \sum_{\substack{m \leq t^\alpha \\ m \in E_0 \\ P(m) \leq t}} \frac{a(m)}{m}. \tag{2.9}$$

Let H be a large constant and observe that

$$\sum_{\substack{a(m) > H \\ P(m) < t}} \frac{a(m)}{m} < \frac{1}{H} \sum_{P(m) < t} \frac{a^2(m)}{m} < \frac{1}{H} \prod_{p < t} \left(1 + \frac{a^2(p)}{p}\right) \leq \frac{c_1}{H} \log t.$$

Thus

$$\sum\nolimits_1 \leq \frac{2c^* c_1}{H} + \sum\nolimits_3,$$

where

$$\sum\nolimits_3 \leq \frac{2Hc^*}{(\log t)} \sum_{\substack{m \leq t^\alpha \\ m \in E_0 \\ P(m) \leq t}} \frac{1}{m}.$$

From (2.8) we obtain that,

$$\sum\nolimits_3 \leq \frac{2Hc^*}{\log t} (\log\log t + \rho(\log\log t)(\alpha \log t))$$

consequently $\sum_3 \to 0$ as $x \to \infty$.

Hence we deduce that

$$\limsup_{x \to \infty} \frac{T(x)}{\mathrm{li}\, x} \leq 2c(\alpha) + \frac{2c^* c_1}{H}. \tag{2.10}$$

Since H and α are arbitrary, $c(\alpha) \to 0 \quad (\alpha \to \infty)$ we obtain that the left hand side of (2.10) is zero.

The proof is completed. ∎

3. Further remarks

We would be able to prove the following statements.

1) There exists a sequence $\rho_x \nearrow \infty$, such that the number of those primes $p \leq x$ for which $p + l$ has a couple of divisors d_1, d_2, such that $\rho_x < d_1 < d_2 < 2d_1$, is $(1 + o(1))\pi(x)$.

2) The same assertion remains true for the unitary divisors instead of the divisors.

3) The relative density of the primes p for which $p + l$ has three divisors $d_1 < d_2 < d_3$ $(< 2d_1)$ exists, and smaller than 1. The same assertion for the whole set of the integers was proved by P. Erdős.

References

[1] H. MAIER and G. TENENBAUM, On the set of divisors of an integer, *Invent. Math.* **76** (1984), 121–128.

[2] R. R. HALL and G. TENENBAUM, *Divisors*, Cambridge Univ. Press, New York, 1998.

[3] K. PRACHAR, *Primzahlverteilung*, Springer-Verlag, Berlin, 1957.

[4] K.-H. INDLEKOFER and N. M. TIMOFEEV, Divisors of shifted primes, *Publ. Math. Debrecen*, **60** (2002), 307–345.

I. Kátai
Department of Computer Algebra,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
and
Applied Number Theory Research Group
Hungarian Academy of Sciences
katai@compalg.inf.elte.hu

M. V. Subbarao
University of Alberta
Edmonton, Alberta
Canada T6G 2G1
m.v.subbarao@ualberta.ca

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 35–46

BIFURCATION IN A PREDATOR–PREY MODEL WITH CROSS DIFFUSION

By

SHABAN ALY and MIKLÓS FARKAS

(Received April 23, 2004)

1. Introduction

In [1] we considered a predator–prey system of Cavani–Farkas type (see [2]) living in a habitat of two identical patches in which the per capita migration rate of each species is influenced only by its own density and we have shown that at a critical value of the bifurcation parameter the system undergoes a Turing bifurcation (see [7]), pattern emerge. In population dynamics there are a lot of problems which are described by a cross-diffusion system (see [3], [4], [5], [6]). In this paper, we consider the case when the per capita migration rate of each species is influenced not only by its own but also by the other one's density, i.e. there is cross diffusion present.

Let $u_1(t,j) :=$ density of prey in patch j at time t and $u_2(t,j) :=$ density of predator in patch j at time t, $j = 1,2$; $t \in R$. The interaction between two species is described as a system of differential equations as follows:

$$
\begin{aligned}
\dot{u}_1(t,1) &= \varepsilon u_1(t,1)\left(1 - \frac{u_1(t,1)}{K}\right) - \frac{\beta u_1(t,1)u_2(t,1)}{\beta + u_1(t,1)}\\
&\quad + d_1(\rho_1(u_2(t,2))u_1(t,2) - \rho_1(u_2(t,1))u_1(t,1)),\\
\dot{u}_2(t,1) &= -\frac{u_2(t,1)(\gamma + \delta u_2(t,1))}{1 + u_2(t,1)} + \frac{\beta u_1(t,1)u_2(t,1)}{\beta + u_1(t,1)}\\
&\quad + d_2(\rho_2(u_1(t,2))u_2(t,2) - \rho_2(u_1(t,1))u_2(t,1)),\\
\dot{u}_1(t,2) &= \varepsilon u_1(t,2)\left(1 - \frac{u_1(t,2)}{K}\right) - \frac{\beta u_1(t,2)u_2(t,2)}{\beta + u_1(t,2)}\\
&\quad + d_1(\rho_1(u_2(t,1))u_1(t,1) - \rho_1(u_2(t,2))u_1(t,2)),
\end{aligned}
\tag{1}
$$

$$\begin{aligned}\dot{u}_2(t,2) = &-\frac{u_2(t,2)(\gamma+\delta u_2(t,2))}{1+u_2(t,2)} + \frac{\beta u_1(t,2)u_2(t,2)}{\beta+u_1(t,2)} \\ &+ d_2(\rho_2(u_1(t,1))u_2(t,1) - \rho_2(u_1(t,2))u_2(t,2)),\end{aligned} \tag{1}$$

where $\varepsilon > 0$ is the specific growth rate of the prey in the absence of predation and without environmental limitation, $\beta > 0$, $K > 0$ are the conversion rate and carrying capacity with respect to the prey, respectively, $\gamma > 0$ and $\delta > 0$ are the minimal mortality and the limiting mortality of the predator, respectively (the natural assumption is $\gamma < \delta$). The meaning of conversion rate or the half saturation constant is that at $u_1 = \beta$ the specific growth rate $\frac{\beta u_1}{\beta+u_1}$ (called also a Holling type functional response) of the predator is equal to half its maximum β. The advantage of the present model over the more often used models is that here the predator mortality is neither a constant nor an unbounded function, still, it is increasing with quantity. $d_i > 0$, $(i = 1,2)$ are the diffusion coefficients and $\rho_1 \in C^1$ is a positive increasing function of u_2, the density of the predator, $\rho_1' > 0$ and $\rho_2 \in C^1$ is a positive decreasing function of u_1 the density of the prey, $\rho_2' < 0$. The idea is that the dependence of the diffusion coefficient on the density of the other species reflects the inclination of a prey (or an activator) to leave a certain patch because of the danger (or the inhibition) and the tendency of a predator (or the inhibition) to stay at a certain patch because of the abundance of prey (or an activator). The functions ρ_i model the cross-diffusion effect. We say that the cross diffusion is strong if $\left|\rho'_{iu_k}\right|$ $(i \neq k)$ is large. If by varying a parameter $\left|\rho'_{iu_k}\right|$ $(i \neq k)$ is increasing then we say that the cross diffusion effect is increasing. If $\rho_i = 1$, $i = 1,2$ then we have mere "self-diffusion".

First we consider the kinetic system without migration, i.e. $d_1 = d_2 = 0$:

$$\begin{aligned}\dot{u}_1(t,1) &= \varepsilon u_1(t,1)\left(1 - \frac{u_1(t,1)}{K}\right) - \frac{\beta u_1(t,1)u_2(t,1)}{\beta+u_1(t,1)}, \\ \dot{u}_2(t,1) &= -\frac{u_2(t,1)(\gamma+\delta u_2(t,1))}{1+u_2(t,1)} + \frac{\beta u_1(t,1)u_2(t,1)}{\beta+u_1(t,1)}, \\ \dot{u}_1(t,2) &= \varepsilon u_1(t,2)\left(1 - \frac{u_1(t,2)}{K}\right) - \frac{\beta u_1(t,2)u_2(t,2)}{\beta+u_1(t,2)}, \\ \dot{u}_2(t,2) &= -\frac{u_2(t,2)(\gamma+\delta u_2(t,2))}{1+u_2(t,2)} + \frac{\beta u_1(t,2)u_2(t,2)}{\beta+u_1(t,2)}.\end{aligned} \tag{2}$$

The following conditions are reasonable and natural:

$$\gamma < \beta \leq \delta, \tag{3}$$

$$\beta < K, \tag{4}$$

$$\gamma < \frac{\beta K}{\beta + K}. \tag{5}$$

Condition (3) ensures that the predator mortality is increasing with density, and that the predator null-cline has a reasonable concave down shape; (4) ensures that for the prey an Allée-effect zone exists where the increase of prey density is favourable to its growth rate; (5) is needed to have a positive equilibrium point of system (2) (see [1]). System (2) is made up by two identical uncoupled systems. Under these conditions each has (the same) positive equilibrium which is the intersection of the null-clines:

$$u_2 = H_1(u_1) := \frac{\varepsilon}{\beta K}(K - u_1)(\beta + u_1), \tag{6}$$

$$u_2 = H_2(u_1) := \frac{(\beta - \gamma)u_1 - \beta\gamma}{(\delta - \beta)u_1 + \beta\delta}. \tag{7}$$

Thus, denoting the coordinates of a positive equilibrium by $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$, these coordinates satisfy $\overline{u}_2 = H_1(\overline{u}_1) = H_2(\overline{u}_1)$.

Note that if $K > \beta$, we have an interval $u_1 \in (0, \frac{K-\beta}{2})$, where the Allée-effect holds, i.e. the increase of the prey quantity is beneficial to its growth rate.

The Jacobian matrix of the system (2) linearized at $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is

$$J_k = \begin{pmatrix} \Theta_1 & -\Theta_2 & 0 & 0 \\ \Theta_3 & -\Theta_4 & 0 & 0 \\ 0 & 0 & \Theta_1 & -\Theta_2 \\ 0 & 0 & \Theta_3 & -\Theta_4 \end{pmatrix}. \tag{8}$$

The characteristic polynomial is

$$D_4(\lambda) = (D_2(\lambda))^2, D_2(\lambda) = \lambda^2 + \lambda(\Theta_4 - \Theta_1) + \Theta_2\Theta_3 - \Theta_1\Theta_4, \tag{9}$$

where

$$\Theta_1 = \frac{\varepsilon \overline{u}_1(K - \beta - 2\overline{u}_1)}{K(\beta + \overline{u}_1)}, \qquad \Theta_2 = \frac{\beta \overline{u}_1}{\beta + \overline{u}_1}, \tag{10}$$

$$\Theta_3 = \frac{\beta^2 \overline{u}_2}{(\beta + \overline{u}_1)^2}, \qquad \Theta_4 = \frac{(\delta - \gamma)\overline{u}_2}{(1 + \overline{u}_2)^2}.$$

The equilibrium point $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ lies in the Allée-effect zone if

$$H_1\left(\frac{k-\beta}{2}\right) < H_2\left(\frac{k-\beta}{2}\right), \tag{11}$$

i.e.

$$\frac{\varepsilon}{4\beta K}(K+\beta)^2 < -1 + \frac{(\delta-\gamma)K}{\beta^2 - \beta K + \delta K}. \tag{12}$$

Assume that

$$\Theta_4 - \Theta_1 > 0 \text{ and } \Theta_2\Theta_3 - \Theta_1\Theta_4 > 0, \tag{13}$$

then the coexistence equilibrium point $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is linearly asymptotically stable.

2. The model with self-diffusion

Model (1) with self-diffusion (i.e., $\rho_i(u) \equiv 1$, $i = 1, 2$) can be written as follows:

(14)

$$\begin{aligned}
\dot{u}_1(t,1) &= \varepsilon u_1(t,1)\left(1 - \frac{u_1(t,1)}{K}\right) - \frac{\beta u_1(t,1)u_2(t,1)}{\beta + u_1(t,1)} + d_1(u_1(t,2) - u_1(t,1)),\\
\dot{u}_2(t,1) &= -\frac{u_2(t,1)(\gamma + \delta u_2(t,1))}{1 + u_2(t,1)} + \frac{\beta u_1(t,1)u_2(t,1)}{\beta + u_1(t,1)} + d_2(u_2(t,2) - u_2(t,1)),\\
\dot{u}_1(t,2) &= \varepsilon u_1(t,2)\left(1 - \frac{u_1(t,2)}{K}\right) - \frac{\beta u_1(t,2)u_2(t,2)}{\beta + u_1(t,2)} + d_1(u_1(t,1) - u_1(t,2)),\\
\dot{u}_2(t,2) &= -\frac{u_2(t,2)(\gamma + \delta u_2(t,2))}{1 + u_2(t,2)} + \frac{\beta u_1(t,2)u_2(t,2)}{\beta + u_1(t,2)} + d_2(u_2(t,1) - u_2(t,2)).
\end{aligned}$$

We see that $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is also a spatially homogeneous equilibrium of the system with self-diffusion. The Jacobian matrix of the system at $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ can be written as:

$$J_D = \begin{pmatrix} \Theta_1 - d_1 & -\Theta_2 & d_1 & 0 \\ \Theta_3 & -\Theta_4 - d_2 & 0 & d_2 \\ d_1 & 0 & \Theta_1 - d_1 & -\Theta_2 \\ 0 & d_2 & \Theta_3 & -\Theta_4 - d_2 \end{pmatrix}, \tag{15}$$

(16)

$$\det(J_D - \lambda I) = \begin{vmatrix} \Theta_1 - d_1 - \lambda & -\Theta_2 & d_1 & 0 \\ \Theta_3 & -\Theta_4 - d_2 - \lambda & 0 & d_2 \\ d_1 & 0 & \Theta_1 - d_1 - \lambda & -\Theta_2 \\ 0 & d_2 & \Theta_3 & -\Theta_4 - d_2 - \lambda \end{vmatrix}.$$

Using the properties of determinant we get

$$\begin{aligned}(17)\qquad &\begin{vmatrix} \Theta_1-\lambda & -\Theta_2 & d_1 & 0 \\ \Theta_3 & -\Theta_4-\lambda & 0 & d_2 \\ 0 & 0 & \Theta_1-2d_1-\lambda & -\Theta_2 \\ 0 & 0 & \Theta_3 & -\Theta_4-2d_2-\lambda \end{vmatrix} \\ &= D_2(\lambda)(\lambda^2+\lambda(\Theta_4-\Theta_1+2(d_1+d_2))+(\Theta_2\Theta_3-\Theta_1\Theta_4)+ \\ &+2d_1\Theta_4-2d_2(\Theta_1-2d_1).\end{aligned}$$

We know that $D_2(\lambda)$ has two roots with negative real parts. By (13), clearly, $\Theta_4-\Theta_1+2(d_1+d_2)>0$. The other polynomial will have a negative and a positive root if the constant term is negative. By the properties of the model and condition (12) the first two terms of the constant are positive. If (12) holds and the parameters have been chosen so that

$$(18)\qquad \Theta_1-2d_1>0.$$

we may increase d_2 and the constant term becomes negative, i.e. the equilibrium $(\overline{u}_1,\overline{u}_2,\overline{u}_1,\overline{u}_2)$ becomes diffusively unstable. The calculations lead to the following Theorem.

THEOREM 2.1. *Under conditions (12), (13), (18) if*

$$(19)\qquad d_2>d_{2crit}=\frac{(\Theta_2\Theta_3-\Theta_1\Theta_4+2d_1\Theta_4)}{2(\Theta_1-2d_1)},$$

then Turing instability occurs.

REMARK 2.2. If (12) and (13) hold and the parameters have been chosen so that

$$(20)\qquad \Theta_1-2d_1<0,$$

then self-diffusion never destabilizes the equilibrium $(\overline{u}_1,\overline{u}_2,\overline{u}_1,\overline{u}_2)$ which is asymptotically stable for the kinetic system, i.e. the equilibrium $(\overline{u}_1,\overline{u}_2,\overline{u}_1,\overline{u}_2)$ is diffusively stable for all values of d_2.

3. The model with cross-diffusion

For model (1) with cross-diffusion response (i.e., $\frac{\partial \rho_i(u)}{\partial u_j} \neq 0, i \neq j$) we see that $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is also a spatially homogeneous equilibrium of the system with cross-diffusion.

The Jacobian matrix of the system with cross-diffusion at $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ can be written as:

(21)
$$J_D = \begin{pmatrix} \Theta_1 - d_1\rho_1 & -\Theta_2 - d_1\rho_1'\overline{u}_1 & d_1\rho_1 & d_1\rho_1'\overline{u}_1 \\ \Theta_3 - d_2\rho_2'\overline{u}_2 & -\Theta_4 - d_2\rho_2 & d_2\rho_2'\overline{u}_2 & d_2\rho_2 \\ d_1\rho_1 & d_1\rho_1'\overline{u}_1 & \Theta_1 - d_1\rho_1 & -\Theta_2 - d_1\rho_1'\overline{u}_1 \\ d_2\rho_2'\overline{u}_2 & d_2\rho_2 & \Theta_3 - d_2\rho_2'\overline{u}_2 & -\Theta_4 - d_2\rho_2 \end{pmatrix},$$

where ρ_1 and ρ_1' are to be taken at $\overline{u}_2$ and ρ_2, ρ_2' at $\overline{u}_1$.

THEOREM 3.1. *Under conditions (12), (13) if*

$$\Theta_1 - 2d_1\rho_1 > 0 \tag{22}$$

and $\rho_2(\overline{u}_1)$ is sufficiently large then Turing instability occurs.

PROOF. $\det(J_D - \lambda I) =$

(23)
$$\begin{vmatrix} \Theta_1 - d_1\rho_1 - \lambda & -\Theta_2 - d_1\rho_1'\overline{u}_1 & d_1\rho_1 & d_1\rho_1'\overline{u}_1 \\ \Theta_3 - d_2\rho_2'\overline{u}_2 & -\Theta_4 - d_2\rho_2 - \lambda & d_2\rho_2'\overline{u}_2 & d_2\rho_2 \\ d_1\rho_1 & d_1\rho_1'\overline{u}_1 & \Theta_1 - d_1\rho_1 - \lambda & -\Theta_2 - d_1\rho_1'\overline{u}_1 \\ d_2\rho_2'\overline{u}_2 & d_2\rho_2 & \Theta_3 - d_2\rho_2'\overline{u}_2 & -\Theta_4 - d_2\rho_2 - \lambda \end{vmatrix}.$$

Using the properties of determinant we get

$$\begin{vmatrix} \Theta_1 - \lambda & -\Theta_2 & d_1\rho_1 & d_1\rho_1'\overline{u}_1 \\ \Theta_3 & -\Theta_4 - \lambda & d_2\rho_2'\overline{u}_2 & d_2\rho_2 \\ 0 & 0 & \Theta_1 - 2d_1\rho_1 - \lambda & -\Theta_2 - 2d_1\rho_1'\overline{u}_1 \\ 0 & 0 & \Theta_3 - 2d_2\rho_2'\overline{u}_2 & -\Theta_4 - 2d_2\rho_2 - \lambda \end{vmatrix} \tag{24}$$

$$\begin{aligned} &= D_2(\lambda)\{\lambda^2 + \lambda[\Theta_4 - \Theta_1 + 2(d_1\rho_1 + d_2\rho_2)] + \Theta_2\Theta_3 - \Theta_1\Theta_4 \\ &\qquad + 2d_1\Theta_4\rho_1 - 2d_2\rho_2(\Theta_1 - 2d_1\rho_1) + 2d_1\overline{u}_1\Theta_3\rho_1' \\ &\qquad - 2d_2\rho_2'\overline{u}_2(\Theta_2 + 2d_1\rho_1'\overline{u}_1)\}. \end{aligned} \tag{25}$$

We know that $D_2(\lambda)$ has two roots with negative real parts. By (13), clearly, $\Theta_4 - \Theta_1 + 2(d_1\rho_1 + d_2\rho_2) > 0$. The other polynomial will have a negative and a positive root if its constant term is negative. This can be achieved if $\rho_2(\overline{u}_1)$ is increased. ∎

REMARK 3.2. As we have mentioned, if (20) holds and there is no cross diffusion then the equilibrium remains stable for any $d_2 > 0$. Still, (22) may hold, i.e. in this case only the cross diffusion effect may destabilize the equilibrium.

REMARK 3.3. If the parameters have been chosen so that

$$\Theta_1 - 2d_1 > 0 \text{ and } \Theta_1 - 2d_1\rho_1 < 0, \tag{26}$$

then the equilibrium $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ remains asymptotically stable for any $d_2 > 0$ and $\rho_2 > 0$ in the cross diffusion case while, as we have seen, it will undergo a Turing bifurcation in the absence of cross diffusion.

4. Numerical investigations

In this section we illustrate the results by the following example and we are looking for conditions which imply Turing instability (diffusion driven instability).

EXAMPLE. We choose

$$\rho_1(u_2) = \frac{m_1 u_2}{1 + u_2}, \qquad \rho_2(u_1) = m_2 \exp\left(\frac{-u_1}{m_2}\right), \qquad m_1,\ m_2 > 0. \tag{27}$$

If $\beta = 0.1$, $\gamma = 0.01$, $\delta = 0.1055$, $\varepsilon = 1$, $K = 1$. The unique positive equilibrium is $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2) = (0.4486, 3.0250, 0.4486, 3.0250)$. We see that this point is in the Allée-effect zone ($0.4486 < 0.45$) and it is asymptotically stable with respect to the kinetic system (2).

For the self-diffusion system (14) we considered d_2 as a bifurcation parameter. In this case at $d_1 = 0.0001$ we have $d_{2crit} = 2.024478$.

If $d_2 = 1$ (resp. 2.5) then, $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is asymptotically stable (resp. unstable).

For the cross-diffusion system we consider m_2 as a bifurcation parameter. In this case at $d_1 = 1$, $d_2 = 1$, $m_1 = 0.001$ and $m_{2crit} \cong 350.7$, we have four eigenvalues $\lambda_i (i = 1, 2, 3, 4)$ such that $\mathrm{Re}\lambda_i < 0$, $(i = 1, 2, 3)$ and $\lambda_4 = 0$.

If $m_2 < m_{2crit} \Rightarrow \mathrm{Re}\lambda_i < 0$, $(i = 1, 2, 3, 4)$, $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is asymptotically stable.

If $m_2 > m_{2crit} \Rightarrow \mathrm{Re}\lambda_i < 0$, $(i = 1, 2, 3)$ and $\lambda_4 > 0$, $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is unstable.

If $d_1 = 0.0001$, $d_{2crit} = 2.5$ and $m_1 = 100$, then, $(\overline{u}_1, \overline{u}_2, \overline{u}_1, \overline{u}_2)$ is asymptotically stable for all m_2.

In this example $\left|\rho'_{2u_1}(u_1, u_2)\right| = \exp\left(-\frac{u_1}{m_2}\right)$. As we see if m_2 is increased for fixed u_1 this derivative is increasing, i.e. the cross diffusion effect is increasing and the spatially homogeneous equilibrium loses its stability. Numerical calculations show that two new spatially non-constant equilibria emerge (see the Table and the Figures), and these equilibria are asymptotically stable.

5. Conclusions

In the present article our interest is to study a prey–predator system of Cavani–Farkas type in two patches in which the per capita migration rate of each species is influenced not only by its own but also by the other one's density, i.e. there is cross diffusion present. Our main result is that a standard (self-diffusion) system may be either stable or unstable, a cross-diffusion response can stabilize an unstable standard system and destabilize a stable standard system. We conclude that the cross migration response is an important factor that should not be ignored when pattern emerges.

Acknowledgement

Research partially supported by the Hungarian National Foundation for Scientific Research.

References

[1] S. A. Aly and M. Farkas, Bifurcation in a predator–prey model in patchy environment with diffusion, *Nonlinear Analysis: Real World Applications* **5** (2004), 519–526.

[2] M. Cavani and M. Farkas, Bifurcation in a predator–prey model with memory and diffusion II: Turing bifurcation; *Acta Math. Hungar.* **63** (1994), 375–393.

[3] M. Farkas, Two ways of modeling cross diffusion, *Nonlinear Analysis, TMA.* **30** (1997), 1225–1233.

[4] M. Farkas, *Dynamical Models in Biology*, Academic Press, 2001.

[5] J. D. Murray, *Mathematical Biology*, Berlin, Springer-Verlag, 1989.

[6] A. OKUBO and S. A. LEVIN, *Diffusion and Ecological Problems: Modern Perspectives*; Second Edition, Springer, Berlin, 2000.

[7] A. M. TURING, The chemical basis of morphogenesis, *Philos. Trans. Roy. Soc. London* **B237** (1953), 37–72, reprinted: *Bull. Math. Biol.* **52** (1990), 153–197.

Shaban Aly

Current Address:
Institute of Mathematics,
Budapest University of Technology,
Budapest, Hungary, H-1521.
shaban@math.bme.hu

Permanent Address:
Department of Mathematics,
Faculty of Science,
Al-Azhar Uni.,
Assiut 71511, EGYPT.
shhaly12@yahoo.com

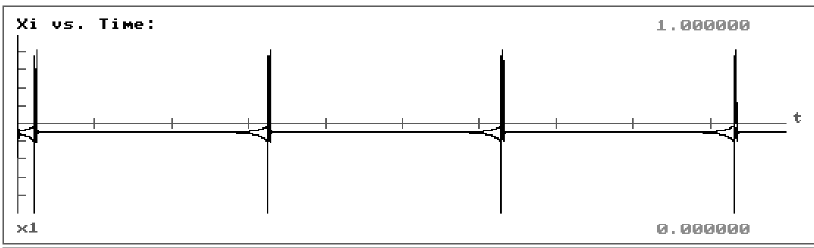
Miklós Farkas

Institute of Mathematics
Budapest University of Technology
H-1521 Budapest
P.O.B. 91.
Hungary
fm@math.bme.hu

TABLE. Equilibria of the Example before and after bifurcation.

m_2	$u_1(t,1)$	$u_2(t,1)$	$u_1(t,2)$	$u_2(t,2)$
350	.4486421535	3.024981563	.4486421535	3.024981563
	.4378285520	3.023718369	.4594667816	3.023905740
355	.4486421535	3.024981563	.4486421535	3.024981563
	.4594667816	3.023905740	.4378285520	3.023718369
	.4293426859	3.021090741	.4679770577	3.021415850
365	.4486421535	3.024981563	.4486421535	3.024981563
	.4679770577	3.021415850	.4293426859	3.021090741
	.4239447856	3.018670501	.4733980189	3.019075240
375	.4486421535	3.024981563	.4486421535	3.024981563
	.4733980189	3.019075240	.4239447856	3.018670501
	.4198086580	3.016422000	.4775559683	3.016882011
385	.4486421535	3.024981563	.4486421535	3.024981563
	.4775559683	3.016882011	.4198086580	3.016422000
	.4149242890	3.013326528	.4824709584	3.013843903
400	.4486421535	3.024981563	.4486421535	3.024981563
	.4824709584	3.013843903	.4149242890	3.013326528

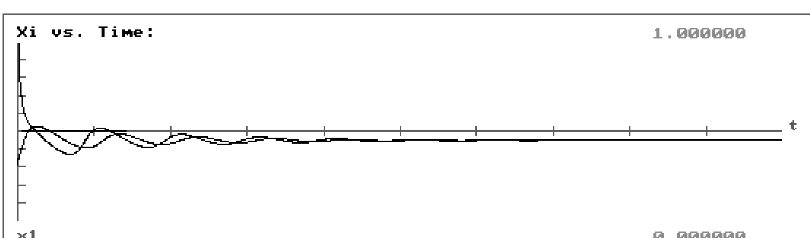
FIGURE 1. Graphs of the coordinate $u_1(t, 1)$ of two solutions of the Example corresponding to the respective initial conditions (0.33, 2.85, 0.5, 2.91), (3.332, 2.88, 0.542, 2.85); (a) for self-diffusion at $d_1 = 0.0001$, $d_2 = 2.5$, (b) for cross-diffusion at $d_1 = 0.0001$, $d_2 = 2.5$, $m_1 = 100$ and $m_2 = 1$, (Figure produced by applying PHASER).



LAST Xi:

IC	Time	X	Y	Z
1	199998.14	0.448642	3.024982	0.448642
2	199998.14	0.448642	3.024982	0.448642
1	199998.69	0.448642	3.024982	0.448642
2	199998.69	0.448642	3.024982	0.448642
1	199999.24	0.448642	3.024982	0.448642
2	199999.24	0.448642	3.024982	0.448642
1	199999.79	0.448642	3.024982	0.448642
2	199999.79	0.448642	3.024982	0.448642

(a)

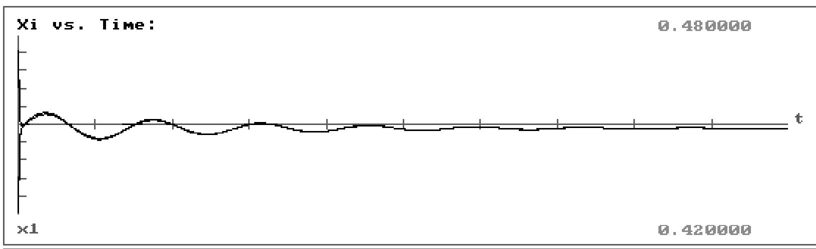


LAST Xi:

IC	Time	X	Y	Z
1	698.50000	0.448430	3.024739	0.448430
2	698.50000	0.448792	3.024567	0.448792
1	699.00000	0.448439	3.024731	0.448439
2	699.00000	0.448809	3.024578	0.448809
1	699.50000	0.448450	3.024723	0.448450
2	699.50000	0.448825	3.024591	0.448825
1	700.00000	0.448460	3.024716	0.448460
2	700.00000	0.448841	3.024604	0.448841

(b)

FIGURE 2. Graphs of the coordinate $u_1(t, 1)$ of two solutions of the Example corresponding to the respective initial conditions (0.423, 3.018, 0.473, 3.02), (0.4733, 3.018, 0.423, 3.0186); (a) for self-diffusion at $d_1 = 1$, $d_2 = 1$, (b) for cross-diffusion at $d_1 = 1$, $d_2 = 1$, $m_1 = 0.001$ and $m_2 = 375$, (Figure produced by applying PHASER).



LAST Xi:

IC	Time	X	Y	Z
1	498.30000	0.448717	3.024905	0.448717
2	498.30000	0.448728	3.024900	0.448728
1	498.85000	0.448720	3.024910	0.448720
2	498.85000	0.448732	3.024906	0.448732
1	499.40000	0.448723	3.024915	0.448723
2	499.40000	0.448735	3.024912	0.448735
1	499.95000	0.448726	3.024920	0.448726
2	499.95000	0.448738	3.024918	0.448738

(a)

LAST Xi:

IC	Time	X	Y	Z
1	999.99400	0.423511	3.018450	0.473834
2	999.99400	0.473667	3.018944	0.423677
1	999.99600	0.423511	3.018450	0.473834
2	999.99600	0.473667	3.018944	0.423677
1	999.99800	0.423511	3.018450	0.473834
2	999.99800	0.473667	3.018944	0.423677
1	1000.0000	0.423511	3.018450	0.473834
2	1000.0000	0.473667	3.018944	0.423677

(b)

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 47–64

RELATIONAL STRUCTURES, MULTIALGEBRAS AND INVERSE LIMITS

By

COSMIN PELEA

(Received: August 8, 2004)

1. Introduction

The study of multialgebras (hyperstructures) began more than 65 years ago and they were used in different areas of mathematics and in some applied sciences (see [2]). Multialgebras are particular relational systems which generalize the universal algebras. The starting point of this paper is in [6], [7] and [15]. The papers [6] and [7] present the construction of the inverse limit for particular inverse systems for some cases of (semi)hypergroups. The construction is a generalization for the same construction made for corresponding universal algebras as well as a particularization of the construction mentioned in [4] for relational systems. In [15] is proved that the category of multialgebras (with the homomorphisms obtained from the relational systems homomorphisms) has finite products and equalizers, hence it is finitely complete. With no major changes in the proof it can be shown that the category studied in [15] is a category with products, hence this is a complete category. But the notion of multioperation used in [15] does not exclude the empty set from its range and the proof that these multistructures form a category with equalizers uses this fact. Such a multistructure does not satisfy the characterization theorem given by G. Grätzer in [3]. A natural question is if the subcategory whose objects are the multialgebras characterized by Grätzer is closed under the formation of limits in the category studied in [15]. We will show that the answer is negative by proving that the subcategory of the multialgebras of type $(n_\gamma)_{\gamma<o(\tau)}$ is not closed under the formation of (inverse) limits of inverse systems in the category of the relational systems of type $(n_\gamma+1)_{\gamma<o(\tau)}$. We will also examine the conditions under which the inverse

limit of an inverse system of multialgebras (considered in the category of relational systems) is a multialgebra.

The main results of this paper are presented in Section 4 and Section 5. The first important result in Section 4 is Proposition 4.2. This proposition states that if we consider an inverse system of multialgebras for which the carrier has a cofinal subset for which the resulting inverse system of multialgebras has an inverse limit which is a multialgebra, then the inverse limit of the given inverse system of multialgebras is also a multialgebra and the two inverse limit multialgebras are isomorphic. As we will see, some of the results presented in [2], [6] and [7] can be easily obtained from this proposition. In the same section we will prove that a class of multialgebras closed under the formation of the isomorphic images and under the formation of the inverse limits of inverse systems with well-ordered carriers is closed under the formation of the inverse limits of arbitrary inverse systems.

As it is shown in [14], an important tool in the hyperstructure theory is the fundamental relation of a multialgebra. The factorization of a multialgebra with the fundamental relation gives a universal algebra. As we have seen in [10], this way we obtain a functor from the category of the multialgebras of a given type into the category of the universal algebras of this type. Another purpose of this paper is to find some conditions for this functor to preserve the inverse limits of inverse systems of multialgebras. Some results in this direction will be presented in Section 5.

2. Preliminaries

Let $\tau = (n_\gamma)_{\gamma<o(\tau)}$ be a sequence over $\mathbb{N} = \{0, 1, \ldots\}$, where $o(\tau)$ is an ordinal and for any $\gamma < o(\tau)$, let $\mathbf{f}_\gamma$ be a symbol of an n_γ-ary (multi)operation and let us consider the algebra of the n-ary terms (of type τ) $\mathfrak{P}^{(n)}(\tau) = \left(\mathbf{P}^{(n)}(\tau), (f_\gamma)_{\gamma<o(\tau)}\right)$.

Let A be a set and let $P^*(A)$ be the set of nonempty subsets of A. Let $\mathfrak{A} = (A, (f_\gamma)_{\gamma<o(\tau)})$ be a multialgebra, where, for any $\gamma < o(\tau)$, $f_\gamma : A^{n_\gamma} \to P^*(A)$ is the multioperation of arity n_γ that corresponds to the symbol $\mathbf{f}_\gamma$. An empty set A is admissible if there are no nullary multioperations among the multioperations f_γ, $\gamma < o(\tau)$. Of course, any universal algebra is a multialgebra (we can identify a one element set with its element).

We remind the reader that if $\mathfrak{A} = (A, (f_\gamma)_{\gamma<o(\tau)})$ is a multialgebra and $B \subseteq A$, we say that $\mathfrak{B} = (B, (f_\gamma)_{\gamma<o(\tau)})$ is a submultialgebra of $\mathfrak{A}$ if for any

$\gamma < o(\tau)$ and $b_0, \ldots, b_{n_\gamma - 1} \in B$, we have $f_\gamma(b_0, \ldots, b_{n_\gamma - 1}) \subseteq B$. If $n \in \mathbb{N}$, $\mathbf{p} \in \mathbf{P}^{(n)}(\tau)$ and $b_0, \ldots, b_{n-1} \in B$ then $p(b_0, \ldots, b_{n-1}) \subseteq B$ (see [1]).

As in [13], we can see the multialgebra $\mathfrak{A} = (A, (f_\gamma)_{\gamma < o(\tau)})$ as a relational system $(A, (r_\gamma)_{\gamma < o(\tau)})$ if we consider that, for any $\gamma < o(\tau)$, r_γ is the $n_\gamma + 1$-ary relation defined by

$$(a_0, \ldots, a_{n_\gamma - 1}, a_{n_\gamma}) \in r_\gamma \;\Leftrightarrow\; a_{n_\gamma} \in f_\gamma(a_0, \ldots, a_{n_\gamma - 1}). \tag{1}$$

Of course, if we keep unchanged some of the multioperations which are operations, we can see $\mathfrak{A}$ as a first order structure. It is clear that a submultialgebra of the multialgebra $\mathfrak{A}$ is a substructure of the first order structure obtained this way, but not any substructure of $\mathfrak{A}$ is a submultialgebra. Such an example can be found in Section 4 (but it is not difficult to find easier examples).

We mention that the objects of the categories studied in [15] are obtained by seeing each relation r_γ of a relational system $(A, (r_\gamma)_{\gamma < o(\tau)})$ of type $(n_\gamma + 1)_{\gamma < o(\tau)}$ as a function $A^{n_\gamma} \to P(A)$ by using (1).

If we define for any $\gamma < o(\tau)$ and for any $A_0, \ldots, A_{n_\gamma - 1} \in P^*(A)$

$$f_\gamma(A_0, \ldots, A_{n_\gamma - 1}) = \bigcup \{f_\gamma(a_0, \ldots, a_{n_\gamma - 1}) \mid \; a_i \in A_i, \; i \in \{0, \ldots, n_\gamma - 1\}\},$$

we obtain a universal algebra on $P^*(A)$ (see [11]). We denote this algebra by $\mathfrak{P}^*(A)$. As in [4], we can construct the algebra $\mathfrak{P}^{(n)}(\mathfrak{P}^*(A))$ of the n-ary term functions on $\mathfrak{P}^*(A)$ for any $n \in \mathbb{N}$.

The *fundamental relation* of the multialgebra $\mathfrak{A}$ is the transitive closure α^* of the relation α given on A as follows: for $x, y \in A$, $x \;\alpha\; y$ if and only if

$$x, y \in p(a_0, \ldots, a_{n-1}) \text{ for some } n \in \mathbb{N},\; \mathbf{p} \in \mathbf{P}^{(n)}(\tau) \text{ and } a_0, \ldots, a_{n-1} \in A,$$

where $p \in P^{(n)}(\mathfrak{P}^*(A))$ is the term function induced by $\mathbf{p}$ on $\mathfrak{P}^*(A)$. The relation α^* is the smallest equivalence relation on A with the property that the factor multialgebra $\mathfrak{A}/\alpha^*$ is a universal algebra (see [9] and [10]). The universal algebra $\mathfrak{A}/\alpha^*$ is called the *fundamental algebra* of the multialgebra $\mathfrak{A}$. We denote by $\overline{\mathfrak{A}}$ the fundamental algebra of $\mathfrak{A}$. We also denote by φ_A the canonical projection of $\mathfrak{A}$ onto $\overline{\mathfrak{A}}$ and by $\overline{a}$ the class $\alpha^*\langle a\rangle = \varphi_A(a)$ of the element $a \in A$.

We remind that for an equivalence relation ρ on A we obtain a multialgebra on A/ρ by defining the multioperations in the factor multialgebra $\mathfrak{A}/\rho$ as follows: for any $\gamma < o(\tau)$, $f_\gamma(\rho\langle a_0\rangle, \ldots, \rho\langle a_{n_\gamma - 1}\rangle)$ is the set

$$\{\rho\langle b\rangle \mid b \in f_\gamma(b_0, \ldots, b_{n_\gamma - 1}),\; a_i \rho b_i,\; i \in \{0, \ldots, n_\gamma - 1\}\}$$

($\rho\langle x\rangle$ denotes the class of x modulo ρ (see [3])).

A map $h : A \to B$ between the multialgebras $\mathfrak{A}$ and $\mathfrak{B}$ of type τ is called homomorphism if for any $\gamma < o(\tau)$ and $a_0, \ldots, a_{n_\gamma - 1} \in A$ we have

$$h(f_\gamma(a_0, \ldots, a_{n_\gamma - 1})) \subseteq f_\gamma(h(a_0), \ldots, h(a_{n_\gamma - 1})).$$

It is easy to see that the multialgebra homomorphisms are obtained from the homomorphisms of relational systems using (1). The definition of the multioperations of $\mathfrak{A}/\rho$ allows us to see the canonical map from A to A/ρ as a homomorphism of multialgebras. The map h is called an ideal homomorphism if for any $\gamma < o(\tau)$ and for all $a_0, \ldots, a_{n_\gamma - 1} \in A$ we have

$$h(f_\gamma(a_0, \ldots, a_{n_\gamma - 1})) = f_\gamma(h(a_0), \ldots, h(a_{n_\gamma - 1})).$$

A bijective map h is a multialgebra isomorphism if both h and h^{-1} are multialgebra homomorphisms. As it shown in [11], the multialgebra isomorphisms can be characterized as being those bijective homomorphisms which are ideal.

REMARK 1. From the steps of construction of a term (function) it follows that for a homomorphism $h : A \to B$, if $n \in \mathbb{N}$, $\mathbf{p} \in \mathbf{P}^{(n)}(\tau)$ and $a_0, \ldots, a_{n-1} \in A$ then

$$h(p(a_0, \ldots, a_{n-1})) \subseteq p(h(a_0), \ldots, h(a_{n-1})).$$

If h is an ideal homomorphism then

$$h(p(a_0, \ldots, a_{n-1})) = p(h(a_0), \ldots, h(a_{n-1})).$$

In [10] we proved the following:

THEOREM 2.1. *If $\mathfrak{A}$, $\mathfrak{B}$ are multialgebras and $\overline{\mathfrak{A}}$, $\overline{\mathfrak{B}}$ respectively are their fundamental algebras and if $f : A \to B$ is an ideal homomorphism then there exists only one homomorphism of universal algebras $\overline{f} : \overline{A} \to \overline{B}$ such that the following diagram is commutative:*

(2)
$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\varphi_A"] & B \arrow[d, "\varphi_B"] \\ \overline{A} \arrow[r, "\overline{f}"] & \overline{B} \end{tikzcd} .$$

The proof uses only the fact that f is a homomorphism, so the statement holds if we drop the property of f being ideal.

COROLLARY 2.2. *a) If $\mathfrak{A}$ is a multialgebra then* $\overline{1_A} = 1_{\overline{A}}$.

b) If $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ are multialgebras of the same type τ and if $f : A \to B$, $g : B \to C$ are homomorphisms, then $\overline{g \circ f} = \overline{g} \circ \overline{f}$.

We can easily construct the category of multialgebras of the same type τ where the morphisms are considered to be the homomorphisms and the composition of two morphisms is the usual map composition. It is known that the universal algebras of the same type τ together with the homomorphisms between them form a category which is, obviously, a full subcategory in the category of the multialgebras introduced above. We will denote by $\mathbf{Malg}(\tau)$ the category of the multialgebras of type τ and by $\mathbf{Alg}(\tau)$ the category of the universal algebras of type τ.

REMARK 2. From Corollary 2.2 it results that we can define a functor F from $\mathbf{Malg}(\tau)$ into $\mathbf{Alg}(\tau)$ as follows: $F(\mathfrak{A}) = \overline{\mathfrak{A}}$, for any multialgebra $\mathfrak{A}$ of type τ, and $F(f) = \overline{f}$ which makes diagram (2) commutative for any homomorphism f between the multialgebras $\mathfrak{A}$ and $\mathfrak{B}$ of type τ.

3. Direct products of multialgebras

If we consider a family $(\mathfrak{A}_i \mid i \in I)$ of multialgebras of type τ using (1) and the definition of the direct product of a family of relational systems presented in [4], we can organize the Cartesian product $\prod_{i\in I} A_i$ as a relational system which is a multialgebra $\prod_{i\in I} \mathfrak{A}_i$ of type τ with the multioperations defined as follows:

$$f_\gamma\left((a_i^0)_{i\in I}, \ldots, (a_i^{n_\gamma - 1})_{i\in I}\right) = \prod_{i\in I} f_\gamma\left(a_i^0, \ldots, a_i^{n_\gamma - 1}\right),$$

for any $\gamma < o(\tau)$. This multialgebra is called the *direct product* of the multialgebras $(\mathfrak{A}_i \mid i \in I)$. We observe that the projections of the product, e_i^I, $i \in I$, are multialgebra (ideal) homomorphisms. So, we have:

PROPOSITION 3.1. *The multialgebra $\prod_{i\in I} \mathfrak{A}_i$, with the projections e_i^I, is the product of the multialgebras $(\mathfrak{A}_i \mid i \in I)$ in the category* $\mathbf{Malg}(\tau)$.

4. Inverse limits of inverse systems of multialgebras

Let $(I, \leq)$ be a preordered directed set and let us consider an inverse system $(\mathfrak{A}_i \mid i \in I)$ of multialgebras of type τ together with the homomorphisms $(\varphi_k^j : A_j \to A_k \mid j, k \in I,\ j \geq k)$. We remind the reader that for any $i \in I$, $\varphi_i^i = 1_{A_i}$ and if $j \geq k \geq l$, we have $\varphi_l^k \circ \varphi_k^j = \varphi_l^j$. We consider the inverse limit of the inverse system of sets $(A_i \mid i \in I)$

$$A^\infty = \Big\{(a_i)_{i\in I} \in \prod_{i\in I} A_i \mid \forall j, k \in I,\ j \leq k,\ \varphi_j^k(a_k) = a_j\Big\}.$$

In [4] it is mentioned that the inverse limits for first order structures are defined the same way as for algebras, as suitable substructures of the direct product. If we see each n_γ-ary multioperation in each $\mathfrak{A}_i$ as an $(n_\gamma + 1)$-ary relation r_γ as in (1), we obtain the definitions for the relations on A^∞ : given $\gamma < o(\tau)$ and $\big(a_i^0\big)_{i\in I}, \ldots, \big(a_i^{n_\gamma - 1}\big)_{i\in I}, (a_i)_{i\in I} \in A^\infty$ we have

$$\big(\big(a_i^0\big)_{i\in I}, \ldots, \big(a_i^{n_\gamma-1}\big)_{i\in I}, (a_i)_{i\in I}\big) \in r_\gamma \Leftrightarrow a_i \in f_\gamma\big(a_i^0, \ldots, a_i^{n_\gamma-1}\big),\ \forall i \in I.$$

Since we are dealing with multialgebras our question is whether the relational system obtained in this way is a multialgebra. If the answer were affirmative then, using again (1), it would follow that its multioperations would be defined by:

$$f_\gamma\big(\big(a_i^0\big)_{i\in I}, \ldots, \big(a_i^{n_\gamma-1}\big)_{i\in I}\big) = \prod_{i\in I} f_\gamma\big(a_i^0, \ldots, a_i^{n_\gamma-1}\big) \cap A^\infty, \tag{3}$$

for every $\gamma < o(\tau)$ and $\big(a_i^0\big)_{i\in I}, \ldots, \big(a_i^{n_\gamma-1}\big)_{i\in I} \in A^\infty$.

REMARK 3. Let us remark that the inverse limit of an inverse system of sets $(A_i \mid i \in I)$, A^∞ is not necessarily a submultialgebra of $\prod_{i\in I} \mathfrak{A}_i$, so, the intersection with A^∞ cannot be omitted in (3).

EXAMPLE 1. Let us consider the finite set of positive integers $I = \{1, 2\}$ ordered with the usual relation $\leq$, induced from $\mathbb{N}$, let us consider the hypergroupoids $(H_1, \circ)$, $(H_2, \circ)$ defined on $H_1 = H_2 = \{x, y\}$ by

$$x \circ x = x \circ y = y \circ x = y \circ y = \{x, y\}$$

and let us consider the (ideal) homomorphisms $\varphi_1^1 = 1_{H_1}$, $\varphi_2^2 = 1_{H_2}$ and

$$\varphi_1^2 : H_2 \to H_1,\ \varphi_1^2(x) = y,\ \varphi_1^2(y) = x.$$

Then $A^\infty = \{(x,y),(y,x)\}$ is not a subhypergroupoid in $H_1 \times H_2$.

REMARK 4. The functions f_γ given by (3) are not always multioperations on A^∞. Even if $A^\infty \neq \emptyset$, the intersection in the second member of the equality can be the empty set. As a matter of fact, since for any $\gamma < o(\tau)$, $\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n\gamma-1}\right)_{i\in I} \in A^\infty$ and for any $j,k \in I$, $j \geq k$,

$$\varphi_k^j\left(f_\gamma\left(a_j^0, \ldots, a_j^{n\gamma-1}\right)\right) \subseteq f_\gamma\left(a_k^0, \ldots, a_k^{n\gamma-1}\right),$$

it follows that for a given $\gamma < o(\tau)$, and $\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n\gamma-1}\right)_{i\in I} \in A^\infty$, the family

$$\left(f_\gamma\left(a_i^0, \ldots, a_i^{n\gamma-1}\right) \mid i \in I\right)$$

of sets together with the restrictions of the maps φ_k^j to these sets form an inverse system of sets and the second member in (3) is the inverse limit of this inverse system of sets. So $f_\gamma\left(\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n\gamma-1}\right)_{i\in I}\right)$ can be empty even if A^∞ is not.

EXAMPLE 2. In [5], Higman and Stone present an example of an inverse system of (countable) sets, with surjective maps and empty inverse limit: let ω_1 be the first uncountable ordinal and for $\alpha < \omega_1$,

$$E_\alpha = \{\gamma \mid \gamma \leq \alpha\},\ F_\alpha = \{f \in \mathbb{R}^{E_\alpha} \mid f \text{ is strictly increasing}\};$$

and for $\alpha < \beta < \omega_1$, let

$$\theta_\alpha^\beta : F_\beta \to F_\alpha,\ \theta_\alpha^\beta(f) = f|_{E_\alpha} \text{ (the restriction of } f \text{ to } E_\alpha).$$

The authors define by transfinite induction a family of subsets S_α of F_α for which $|S_\alpha| = \aleph_0$ and $\theta_\alpha^\beta(S_\beta) = S_\alpha$ whenever $\alpha < \beta$, such that the inverse system $(S_\alpha \mid \alpha < \omega_1)$ (with the corresponding restrictions of the functions θ_α^β) has the desired property.

Starting from this example, we will consider for each $1 \leq \alpha < \omega_1$,

$$A_\alpha = S_\alpha \cup \{0_{E_\alpha}\},\ \text{where } 0_{E_\alpha} : E_\alpha \to \mathbb{R},\ 0_{E_\alpha}(\gamma) = 0.$$

We define a hyperproduct $\circ$ on A_α by taking

$$f \circ g = \begin{cases} S_\alpha, & \text{if } f = 0_{E_\alpha} = g \text{ or } f \neq 0_{E_\alpha} \neq g \\ 0_{E_\alpha}, & \text{otherwise.} \end{cases}$$

The maps

$$\varphi_\alpha^\beta : A_\beta \to A_\alpha,\ \varphi_\alpha^\beta(f) = f|_{E_\alpha}\ (\alpha < \beta)$$

are surjective (ideal) homomorphisms. Thus we obtain an inverse system of hypergroupoids.

We observe that $\varphi_\alpha^\beta|_{S_\alpha} = \theta_\alpha^\beta|_{S_\alpha}$ and that A^∞ is not empty since $(0_{E_\alpha})_{1\le\alpha<\omega_1} \in A^\infty$. In fact $A^\infty = \{(0_{E_\alpha})_{1\le\alpha<\omega_1}\}$, for if A^∞ had an element different from this, it would follow that this element belongs to the inverse limit of the inverse system of the sets $(S_\alpha \mid 1 \le \alpha < \omega_1)$, which is impossible, since this inverse limit is empty.

Now it is easy to see that (3) is not always the definition of a hyperproduct on A^∞ since in our case $(0_{E_\alpha})_{1\le\alpha<\omega_1} \circ (0_{E_\alpha})_{1\le\alpha<\omega_1} = \emptyset$.

REMARK 5. In order to obtain a multialgebra $\mathfrak{A}^\infty$ on $A^\infty \neq \emptyset$ as above it would be required that for any $\gamma < o(\tau)$ and $\left(a_i^0\right)_{i\in I}, \dots, \left(a_i^{n_\gamma-1}\right)_{i\in I} \in A^\infty$,

$$f_\gamma\left(\left(a_i^0\right)_{i\in I}, \dots, \left(a_i^{n_\gamma-1}\right)_{i\in I}\right) = \varprojlim_{i\in I} f_\gamma\left(a_i^0, \dots, a_i^{n_\gamma-1}\right) \neq \emptyset$$

hold. As it is shown in [4, §21, Theorem 1], a case when this happens is given by the condition that for every $i \in I$, $\gamma < o(\tau)$ and $a_i^0, \dots, a_i^{n_\gamma-1} \in A_i$, $f_\gamma\left(a_i^0, \dots, a_i^{n_\gamma-1}\right)$ is nonempty and finite.

REMARK 6. This is the case of universal algebras for which the sets $f_\gamma\left(a_i^0, \dots, a_i^{n_\gamma-1}\right)$ are one-element sets. It is also clear that if for some $\gamma < o(\tau)$, f_γ is an operation in all the multialgebras $\mathfrak{A}_i$, then f_γ is an operation in $\mathfrak{A}^\infty$. In this case, the definition of f_γ is the one in [4, §21].

REMARK 7. If $\mathcal{I}$ is the category associated to the preordered set $(I, \le)$, we can see (as in [12]) the inverse system of multialgebras $(\mathfrak{A}_i \mid i \in I)$, with the homomorphisms $(\varphi_k^j \mid j, k \in I,\ j \ge k)$, as a contravariant functor $G : \mathcal{I} \to \mathbf{Malg}(\tau)$.

REMARK 8. If $(A^\infty, (f_\gamma)_{\gamma<o(\tau)})$ is a multialgebra then, for any $j \in I$,

$$\varphi_j^\infty : A^\infty \to A_j,\ \varphi_j^\infty((a_i)_{i\in I}) = a_j$$

is a multialgebra homomorphism.

Indeed, since the map φ_j^∞ is the restriction of e_j^I to A^∞, if we take $\gamma < o(\tau)$, $(a_i^0)_{i\in I}, \dots, (a_i^{n_\gamma - 1})_{i\in I} \in A^\infty$ we have

$$\begin{aligned}
\varphi_j^\infty\left(f_\gamma\left((a_i^0)_{i\in I}, \dots, (a_i^{n_\gamma-1})_{i\in I}\right)\right) &= \varphi_j^\infty\left(\prod_{i\in I} f_\gamma\left(a_i^0, \dots, a_i^{n_\gamma-1}\right) \cap A^\infty\right)\\
&\subseteq e_j^I\left(\prod_{i\in I} f_\gamma\left(a_i^0, \dots, a_i^{n_\gamma-1}\right)\right)\\
&= f_\gamma\left(a_j^0, \dots, a_j^{n_\gamma-1}\right)\\
&= f_\gamma\left(\varphi_j^\infty\left((a_i^0)_{i\in I}\right), \dots, \varphi_j^\infty\left((a_i^{n_\gamma-1})_{i\in I}\right)\right).
\end{aligned}$$

REMARK 9. As it can be easily seen from the previous examples, the problems which appear when asking for the inverse limit of an inverse system of multialgebras to be a multialgebra are not solved if we are dealing with ideal homomorphisms. Moreover, in this situation new problems arise since the maps φ_j^∞ $(j \in I)$ are not always ideal homomorphisms.

To illustrate this, we built up an example starting once again from the example of Higman and Stone, as in Example 2.

EXAMPLE 3. Take the sets $(A_\alpha \mid \alpha < \omega_1)$, the maps φ_α^β as in Example 2 and define on each A_α, the hyperproduct $\circ$ by

$$f \circ g = A_\alpha.$$

Then, using (3), we obtain a hypergroup(oid) on $A^\infty = \{(0_{E_\alpha})_{\alpha<\omega_1}\}$ with the hyperproduct

$$(0_{E_\alpha})_{\alpha<\omega_1} \circ (0_{E_\alpha})_{\alpha<\omega_1} = (0_{E_\alpha})_{\alpha<\omega_1}.$$

The maps φ_α^β are ideal homomorphisms of hypergroupoids and for any $1 \leq \beta < \omega_1$ the map

$$\varphi_\beta^\infty : A^\infty \to A_\beta,\ \varphi_\beta^\infty((0_{E_\alpha})_{\alpha<\omega_1}) = 0_{E_\beta}$$

is not an ideal homomorphism because

$$\begin{aligned}
\varphi_\beta^\infty((0_{E_\alpha})_{\alpha<\omega_1} \circ (0_{E_\alpha})_{\alpha<\omega_1}) = \varphi_\beta^\infty((0_{E_\alpha})_{\alpha<\omega_1}) = 0_{E_\beta} \neq A_\beta = 0_{E_\beta} \circ 0_{E_\beta}\\
= \varphi_\beta^\infty((0_{E_\alpha})_{\alpha<\omega_1}) \circ \varphi_\beta^\infty((0_{E_\alpha})_{\alpha<\omega_1}).
\end{aligned}$$

Note that in Remark 5 we presented a necessary and sufficient condition for a category of multialgebras (of type τ) to be a subcategory of the category of the relational systems (of type $(n_\gamma+1)_{\gamma<o(\tau)}$) closed under the inverse limits of inverse systems. Thus we have:

THEOREM 4.1. *If* $\mathfrak{A}^\infty = (A^\infty, (f_\gamma)_{\gamma<o(\tau)})$ *is a multialgebra then, together with the homomorphisms* $(\varphi_j^\infty \mid j \in I)$*, it is the inverse limit of the functor* G*.*

The first of the following diagrams:

$$\begin{tikzcd} A_k \arrow[r, "\varphi_j^k"] & A_j \\ & A^\infty \arrow[ul, "\varphi_k^\infty"] \arrow[u, "\varphi_j^\infty"'] \end{tikzcd} \qquad \begin{tikzcd} A_k \arrow[r, "\varphi_j^k"] & A_j \\ A' \arrow[u, "\alpha_k"] \arrow[ur, "\alpha_j"'] & \end{tikzcd} \qquad \begin{tikzcd} & A_j \\ A' \arrow[ur, "\alpha_j"] \arrow[r, "\mu"] & A^\infty \arrow[u, "\varphi_i^\infty"'] \end{tikzcd}$$

is commutative and whenever a multialgebra $\mathfrak{A}' = (A', (f_\gamma)_{\gamma<o(\tau)})$, together with a family $(\alpha_j : A' \to A_j \mid j \in I)$ of homomorphisms make the second diagram commutative, there exists a unique homomorphism $\mu : A' \to A^\infty$, given by the equality $\mu(x) = (\alpha_i(x))_{i\in I}$, such that the third diagram is commutative.

The last three results of this section are generalizations for some results presented for universal algebras in [4, §21].

From now on we will consider that $(I, \leq)$ is a directed partially ordered set (unless we will specify something else). Let $\mathcal{A} = (\mathfrak{A}_i \mid i \in I)$ be an inverse system of multialgebras and let us consider $J \subseteq I$ such that $(J, \leq)$ is also a directed partially ordered set. We will denote by $\mathcal{A}_J$ the inverse system of multialgebras $(\mathfrak{A}_i \mid i \in J)$ whose carrier is $(J, \leq)$ and whose homomorphisms are φ_j^i, with $i, j \in J$, $i \leq j$. If J is cofinal with $(I, \leq)$, from the proof of [4, §21, Lemma 7] it results that if we consider the inverse systems of sets $(A_i \mid i \in I)$ and $(A_i \mid i \in J)$, the canonical projection $(a_i)_{i\in I} \mapsto (a_i)_{i\in J}$ furnishes a bijection ψ between $\varprojlim_{i\in I} A_i$ and $\varprojlim_{i\in J} A_i$.

PROPOSITION 4.2. *Let* $\mathcal{A}$ *be an inverse system of multialgebras with the carrier* $(I, \leq)$ *and let us consider* $J \subseteq I$ *such that* $(J, \leq)$ *is a directed partially ordered set cofinal with* $(I, \leq)$*. The inverse limit* $\varprojlim \mathcal{A}$ *is a multialgebra if and only if* $\varprojlim \mathcal{A}_J$ *is a multialgebra. If this happens, the two multialgebras are isomorphic.*

PROOF. Of course, if the relational systems $\varprojlim \mathscr{A}$ and $\varprojlim \mathscr{A}_J$ are multialgebras, then the isomorphism between them follows from the dual [8, Proposition 2.11, Chapter II]. However, while proving that $\varprojlim \mathscr{A}$ is in $\mathbf{Malg}(\tau)$ if and only if $\varprojlim \mathscr{A}_J$ is in $\mathbf{Malg}(\tau)$ we will find the form of this isomorphism, so, we will also present a proof based on [4, §21, Lemma 7].

Clearly, $\varprojlim_{i\in I} A_i = \emptyset$ if and only if $\varprojlim_{i\in J} A_i = \emptyset$. So, $\varprojlim \mathscr{A}$ and $\varprojlim \mathscr{A}_J$ are multialgebras if and only if for any $\gamma < o(\tau)$, $n_\gamma \neq 0$. If this is the case, they are, trivially, isomorphic.

Let us consider that $\varprojlim_{i\in I} A_i \neq \emptyset \neq \varprojlim_{i\in J} A_i$. The limit $\varprojlim \mathscr{A}$ is a multialgebra whenever for each $\gamma < o(\tau)$ and for all $\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n_\gamma - 1}\right)_{i\in I} \in A^\infty$, the right member of (3) is not the empty set. According to Remark 4, this happens exactly when the inverse system of the nonempty sets $(f_\gamma(a_i^0, \ldots, a_i^{n_\gamma - 1}) \mid i \in I)$, together with the restrictions of the maps φ_k^j, $j \geq k$, to these sets, is not empty. It is easy to see from the definition of ψ that its restrictions

$$\varprojlim_{i\in I} f_\gamma\left(a_i^0, \ldots, a_i^{n_\gamma - 1}\right) \to \varprojlim_{i\in J} f_\gamma\left(a_i^0, \ldots, a_i^{n_\gamma - 1}\right)$$

are also bijective.

As for the multialgebra isomorphism between $\mathfrak{A}^\infty$ and $\mathfrak{A}'^\infty = \varprojlim \mathscr{A}_J$, let us verify that ψ is an ideal homomorphism of multialgebras.

Given $\gamma < o(\tau)$ and $\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n_\gamma - 1}\right)_{i\in I} \in A^\infty$ we have

$$\begin{aligned}
\psi\left(f_\gamma\left(\left(a_i^0\right)_{i\in I}, \ldots, \left(a_i^{n_\gamma - 1}\right)_{i\in I}\right)\right) &= \psi\left(\varprojlim_{i\in I} f_\gamma\left(a_i^0, \ldots, a_i^{n_\gamma - 1}\right)\right) \\
&= \varprojlim_{i\in J} f_\gamma\left(a_i^0, \ldots, a_i^{n_\gamma - 1}\right) \\
&= f_\gamma\left(\left(a_i^0\right)_{i\in J}, \ldots, \left(a_i^{n_\gamma - 1}\right)_{i\in J}\right) \\
&= f_\gamma\left(\psi\left(\left(a_i^0\right)_{i\in I}\right), \ldots, \psi\left(\left(a_i^{n_\gamma - 1}\right)_{i\in I}\right)\right),
\end{aligned}$$

and the proof is finished. ∎

REMARK 10. Using this proposition, the constructions of inverse limits from [2], [6] and [7], which are made for inverse systems of (particular) multialgebras with $(I, \leq)$ directed ordered set which has a maximum, are isomorphic to the member of the system having this maximum as an index.

It is clear that such an inverse limit exists and it has all the properties of this member.

Let us consider that the support set I of the carrier $(I,\leq)$ of the inverse system $\mathcal{A}=(\mathfrak{A}_i \mid i\in I)$ of multialgebras can be written as $I=\bigcup_{p\in P} I_p$, where $(I_p,\leq)$ is a directed partially ordered subset of $(I,\leq)$ for each $p\in P$ and $(P,\leq)$ is also a directed partially ordered set such that $I_p\subseteq I_q$, whenever $p,q\in P$, $p\leq q$. We will denote

$$\varprojlim \mathcal{A}=\mathfrak{A}^{\infty}=(A^{\infty},(f_\gamma)_{\gamma<o(\tau)}),\ \varprojlim \mathcal{A}_{I_p}=\mathfrak{A}_p^{\infty}=(A_p^{\infty},(f_\gamma)_{\gamma<o(\tau)})\ (p\in P).$$

For any $p,q\in P,\ p\leq q$ we can define the map

$$\psi_p^q : A_q^{\infty}\to A_p^{\infty},\ \psi_p^q((a_i)_{i\in I_q})=(a_i)_{i\in I_p}.$$

In this way we obtain an inverse system of sets consisting of $(P,\leq)$, the sets A_p^{∞}, and the maps ψ_p^q. We will denote it by $\mathcal{A}/P$.

THEOREM 4.3. *Let us consider that all $\mathfrak{A}_p^{\infty}$, $p\in P$, are multialgebras. Then $\mathcal{A}/P$ is an inverse system of multialgebras and $\varprojlim \mathcal{A}$ is a multialgebra if and only if $\varprojlim \mathcal{A}/P$ is a multialgebra. In this case these multialgebras are isomorphic.*

PROOF. First we will show that for any $p,q\in P$, $p\leq q$, the map ψ_p^q is a multialgebra homomorphism. If we take $\gamma<o(\tau)$, $\left(a_i^0\right)_{i\in I_q}$, …, $\left(a_i^{n_\gamma-1}\right)_{i\in I_q}\in A_q^{\infty}$ then we have

$$\begin{aligned}\psi_p^q(f_\gamma\left(\left(a_i^0\right)_{i\in I_q},\dots,\left(a_i^{n_\gamma-1}\right)_{i\in I_q}\right)) &= \psi_p^q\left(\prod_{i\in I_q} f_\gamma\left(a_i^0,\dots,a_i^{n_\gamma-1}\right)\cap A_q^{\infty}\right)\\ &\subseteq \prod_{i\in I_p} f_\gamma\left(a_i^0,\dots,a_i^{n_\gamma-1}\right)\cap A_p^{\infty}\\ &= f_\gamma\left(\left(a_i^0\right)_{i\in I_p},\dots,\left(a_i^{n_\gamma-1}\right)_{i\in I_p}\right)\\ &= f_\gamma\left(\psi_p^q\left(\left(a_i^0\right)_{i\in I_q}\right),\dots,\psi_p^q\left(\left(a_i^{n_\gamma-1}\right)_{i\in I_q}\right)\right).\end{aligned}$$

We denote by A'^{∞} the inverse limit of the inverse system of sets $\mathcal{A}/P$ and we consider the map

$$\psi : A^{\infty}\to A'^{\infty},\ \psi(g)=h_g,$$

where $h_g \in \prod_{p\in P} A_p^\infty$ is defined by

$$h_g(p) = g|_{I_p} \in A_p^\infty.$$

By [4, §21, Theorem 3] ψ is a bijective map.

Note that $A'^\infty = \emptyset$ if and only if $A^\infty = \emptyset$. Thus, if $A_p^\infty = \emptyset$ for some p then $A'^\infty = A^\infty = \emptyset$. However, if $A'^\infty = A^\infty = \emptyset$ then $\varprojlim \mathcal{A}$ and $\varprojlim \mathcal{A}/P$ are multialgebras if and only if $n_\gamma \neq 0$ for any $\gamma < o(\tau)$, and the two multialgebras are trivially isomorphic in this case. So, we can assume that the sets A_p^∞, A'^∞, A^∞ are not empty.

The inverse limit $\varprojlim \mathcal{A} = \mathfrak{A}^\infty$ is a multialgebra if and only if for any $\gamma < o(\tau)$ and for all $g_0, \ldots, g_{n_\gamma - 1} \in A^\infty$ the inverse limit of the inverse system of sets $(f_\gamma(g_0(i), \ldots, g_{n_\gamma-1}(i)) \mid i \in I)$ is nonempty. But the restriction of ψ to this inverse limit of sets is a bijection between

$$f_\gamma\left(g_0, \ldots, g_{n_\gamma-1}\right) = \varprojlim_{i\in I} f_\gamma(g_0(i), \ldots, g_{n_\gamma-1}(i))$$

and the inverse limit

$$\begin{aligned}\varprojlim_{p\in P}(\varprojlim_{i\in I_p} f_\gamma(g_0(i), \ldots, g_{n_\gamma-1}(i))) &= \varprojlim_{p\in P} f_\gamma(g_0|_{I_p}, \ldots, g_{n_\gamma-1}|_{I_p})\\ &= \varprojlim_{p\in P} f_\gamma(h_{g_0}(p), \ldots, h_{g_{n_\gamma-1}}(p))\\ &= f_\gamma(h_{g_0}, \ldots, h_{g_{n_\gamma-1}}).\end{aligned}$$

Since

$$\begin{aligned}f_\gamma(g_0|_{I_p}, \ldots, g_{n_\gamma-1}|_{I_p}) &= \varprojlim_{i\in I_p} f_\gamma(g_0|_{I_p}(i), \ldots, g_{n_\gamma-1}|_{I_p}(i))\\ &= \varprojlim_{i\in I_p} f_\gamma(g_0(i), \ldots, g_{n_\gamma-1}(i)) \neq \emptyset\end{aligned}$$

it follows that f_γ is a multioperation on A^∞ if and only if f_γ is a multioperation on A'^∞. Moreover, it follows that

$$\psi(f_\gamma(g_0, \ldots, g_{n_\gamma-1})) = f_\gamma(h_{g_0}, \ldots, h_{g_{n_\gamma-1}}) = f_\gamma(\psi(g_0), \ldots, \psi(g_{n_\gamma-1}))).$$

Thus ψ is a multialgebra isomorphism. ∎

As in [4], we will use the term of algebraic class for those classes of multialgebras which are closed under the formation of isomorphic images. We will say that a class of multialgebras is closed under the formation of inverse limits of inverse systems if for any inverse system of multialgebras from this class the inverse limit is a multialgebra from this class. Now we can prove the following theorem:

THEOREM 4.4. *If K is an algebraic class of multialgebras then K is closed under the formation of inverse limits of arbitrary inverse systems if and only if K is closed under the formation of inverse limits of well-ordered inverse systems.*

PROOF. The proof goes as in [4, §21, Theorem 4]. Assume that K is closed under the formation of inverse limits of well-ordered inverse systems. We take an inverse system $\mathcal{A}$ of multialgebras with carrier $(I, \leq)$. If the theorem were not true, then we could choose $\mathfrak{A}$ such that $|I| = \mathfrak{m}$ is the smallest possible with the property that either the inverse limit of $\mathcal{A}$ is not a multialgebra or $\varprojlim \mathcal{A} \notin K$. From Proposition 4.2 we have $\mathfrak{m} \geq \aleph_0$ and using [4, Exercise 44, pp. 73] it follows that we can write $I = \bigcup_{\delta < \alpha} I_\delta$, where α is an ordinal, $(I_\delta, \leq)$ is directed, $I_{\delta_1} \subseteq I_{\delta_2}$ if $\delta_1 \leq \delta_2 < \alpha$, and $|I_\delta| < |I| = \mathfrak{m}$. From the minimality of $\mathfrak{m}$ it follows that $\varprojlim \mathfrak{A}_{I_\delta}$, $\delta < \alpha$, are multialgebras from K, according to our assumption it follows that $\varprojlim \mathcal{A}/P$ is also a multialgebra from K and by Theorem 4.3 $\varprojlim \mathcal{A}$ is a multialgebra from K, a contradiction. ∎

5. On the fundamental algebra of an inverse limit of multialgebras

In general, the fundamental algebra of a product of multialgebras is not the product of their fundamental algebras, as it is shown by the following:

EXAMPLE 4. Let us consider the hypergroupoid $(\mathbb{Z}, \circ)$, where $\mathbb{Z}$ is the set of the integers and the multioperation $\circ$ is defined by

$$x \circ y = \{x + y, x + y + 1\}$$

for any $x, y \in \mathbb{Z}$. It is not difficult to prove that for any $n \in \mathbb{N}^*$, $(\mathbb{Z}^n, \circ)$ is a hypergroup with the fundamental relation $\beta = \mathbb{Z}^n \times \mathbb{Z}^n$. It means that for any $n \in \mathbb{N}^*$ the fundamental group of $(\mathbb{Z}^n, \circ)$ is a one-element group. Now let us consider the product $(\mathbb{Z}^{\mathbb{N}}, \circ)$. The fundamental group of this hypergroup has more than one element. Indeed, $f, g : \mathbb{N} \to \mathbb{Z}$, $f(n) = 0$, $g(n) = n + 1$ $(n \in \mathbb{N})$ are not in the same equivalence class of the fundamental relation of the hypergroup $(\mathbb{Z}^{\mathbb{N}}, \circ)$.

EXAMPLE 5. An useful example of inverse limit of multialgebras can be given in a similar way as Grätzer did in [4] for universal algebras. So, let us consider a set I and a family $(\mathfrak{A}_i \mid i \in I)$ of multialgebras of type τ. We can get an inverse system of multialgebras taking $(J, \subseteq)$ to be the set of all the finite nonempty subsets of the set I, ordered with the set inclusion, $\mathfrak{B}_j = \prod_{i \in j} \mathfrak{A}_i$, for any $j \in J$, and the canonical projections $\varphi_{j_1}^{j_0}$ from $\prod_{i \in j_0} \mathfrak{A}_i$ onto $\prod_{i \in j_1} \mathfrak{A}_i$, for any $j_0 \supseteq j_1$ from J. The inverse limit of this inverse system of multialgebras exists and it is isomorphic to $\prod_{i \in I} \mathfrak{A}_i$.

REMARK 11. Using the previous examples we can obtain an example to show that the functor F from Remark 2 does not preserve the arbitrary inverse limits of inverse systems of multialgebras (hypergroupoids, in this case) even if they exist.

Now we will find a condition for this functor F to preserve the inverse limits of inverse systems of multialgebras from an algebraic class K closed under the formation of the inverse limits of inverse systems.

Let $\mathcal{A} = (\mathfrak{A}_i \mid i \in I)$ be an inverse system of multialgebras with the homomorphisms $(\varphi_j^i \mid i,j \in I,\ i \geq j)$. We will denote by $\overline{\mathcal{A}}$ the inverse system of the fundamental algebras $(\overline{\mathfrak{A}_i} \mid i \in I)$ of the multialgebras from $\mathcal{A}$, with the homomorphisms $(\overline{\varphi_j^i} \mid i,j \in I,\ i \geq j)$. So, if we see the inverse system $\mathcal{A}$ as the functor G from Remark 7 then $\overline{\mathcal{A}}$ is the functor FG.

In this section, we will refer to the inverse limit $\varprojlim \mathcal{A}$ of the inverse system $\mathcal{A}$ as the inverse limit $(\mathfrak{A}^\infty, (\varphi_i^\infty \mid i \in I))$ of G. Clearly, $\varprojlim \overline{\mathcal{A}} = \varprojlim (FG)$. If we denote $(\overline{\mathfrak{A}^\infty}, (\overline{\varphi_i^\infty} \mid i \in I))$ by $\overline{\varprojlim \mathcal{A}}$ we can state the following:

LEMMA 5.1. *Let $\mathcal{A}$ be an inverse system of multialgebras with the carrier $(I, \leq)$ and let us consider $J \subseteq I$ with $(J, \leq)$ a directed partially ordered set cofinal with $(I, \leq)$. Assume that $\varprojlim \mathcal{A}_J$ is a multialgebra. Under these conditions, $\overline{\varprojlim \mathcal{A}}$ is the inverse limit of the inverse system of universal algebras $\overline{\mathcal{A}}$ if and only if $\overline{\varprojlim \mathcal{A}_J}$ is the inverse limit of the inverse system of universal algebras $\overline{\mathcal{A}_J}$.*

PROOF. From the fact that $\varprojlim \mathcal{A}_J$ is a multialgebra it follows that $\varprojlim \mathcal{A}$ is a multialgebra, and, since they are isomorphic (Proposition 4.2), the functor F induces an isomorphism from $\overline{\varprojlim \mathcal{A}}$ onto $\overline{\varprojlim \mathcal{A}_J}$. Of course, the inverse system $\mathscr{A}_J$ of multialgebras determines an inverse system of universal algebras $\overline{\mathscr{A}_J} = \overline{\mathscr{A}}_J$ and the inverse limit algebras $\varprojlim \overline{\mathscr{A}}$ and $\varprojlim \overline{\mathscr{A}}_J$ are isomorphic. From the universal property that characterizes the inverse limit we obtain the homomorphisms:

$$\overline{\varprojlim \mathscr{A}} \to \varprojlim \overline{\mathscr{A}},\ \overline{(a_i)_{i\in I}} \mapsto (\overline{a_i})_{i\in I};$$

$$\overline{\varprojlim \mathscr{A}_J} \to \varprojlim \overline{\mathscr{A}}_J,\ \overline{(a_i)_{i\in J}} \mapsto (\overline{a_i})_{i\in J},$$

which, together with the above mentioned isomorphisms, make the following diagram commutative:

$$\begin{tikzcd}
\overline{\varprojlim \mathscr{A}} \arrow[r] \arrow[d] & \overline{\varprojlim \mathscr{A}_J} \arrow[d] \\
\varprojlim \overline{\mathscr{A}} \arrow[r] & \varprojlim \overline{\mathscr{A}_J}
\end{tikzcd}.$$

Since each one of the conditions for which we want to prove the equivalence makes one of the vertical morphisms an isomorphism, the conclusion of the lemma is immediate. ∎

LEMMA 5.2. *Using the notations from the previous section, let us consider that $\mathfrak{A}_p^{\infty}$, $p \in P$, and $\mathfrak{A}^{\infty}$ are multialgebras. Let us also assume that for each $p \in P$, $\overline{\varprojlim \mathscr{A}_{I_p}}$ is the inverse limit of the inverse system $\overline{\mathscr{A}_{I_p}}$ of universal algebras.*

The universal algebra $\overline{\varprojlim \mathscr{A}}$ is the inverse limit of $\overline{\mathscr{A}}$ if and only if $\overline{\varprojlim \mathscr{A}/P}$ is the inverse limit of $\overline{\mathscr{A}/P}$.

PROOF. Using the fact that for any $p \in P$, $\overline{\varprojlim \mathscr{A}_{I_p}} = \varprojlim \overline{\mathscr{A}_{I_p}}$ it follows that $\overline{\mathscr{A}}/P = \overline{\mathscr{A}/P}$. From [4, §21, Theorem 3] it results an isomorphism from $\varprojlim \overline{\mathscr{A}}$ onto

$$\varprojlim \overline{\mathscr{A}}/P = \varprojlim \overline{\mathscr{A}/P}.$$

Using Theorem 4.3 and the universal property of the inverse limit, we obtain the following commutative diagram:

$$\begin{tikzcd}
\overline{\varprojlim \mathscr{A}} \arrow[r] \arrow[d] & \overline{\varprojlim \mathscr{A}/P} \arrow[d] \\
\varprojlim \overline{\mathscr{A}} \arrow[r] & \varprojlim \overline{\mathscr{A}/P}
\end{tikzcd}.$$

The conclusion follows as in the previous lemma. ∎

If K is a class of multialgebras of type τ then we can obtain a subcategory $\mathscr{K}$ of $\mathbf{Malg}(\tau)$ if we consider morphisms only those homomorphisms which are defined between two multialgebras from K. Consider the composition FU of the functor F with the inclusion functor $U : \mathscr{K} \longrightarrow \mathbf{Malg}(\tau)$. Knowing the definition of U, in the next theorem we may use F instead of FU.

THEOREM 5.3. *Let K be an algebraic class of multialgebras closed under the formation of inverse limits of well-ordered inverse systems. Then F preserves the inverse limits of arbitrary inverse systems of multialgebras from K if and only if F preserves the inverse limits of well-ordered inverse systems of multialgebras from K.*

PROOF. Clearly, $\mathscr{K}$ is a subcategory of the category of relational systems of type $(n_\gamma + 1)_{\gamma < o(\tau)}$ closed under the inverse limits of inverse systems. Let us assume that F preserves the inverse limits of well-ordered inverse systems of multialgebras from K. If F would not preserve the inverse limits of arbitrary inverse systems of multialgebras from K then we could choose an inverse system $\mathscr{A}$ of multialgebras from K with the carrier $(I, \leq)$ such that $\overline{\varprojlim \mathscr{A}}$ is not $\varprojlim \overline{\mathscr{A}}$ and $|I| = \mathfrak{m}$ is the smallest possible with this property. We will continue as in the proof of Theorem 4.4, so we will use the same notations as there. First we obtain (using Lemma 5.1) that $\mathfrak{m} \geq \aleph_0$ and then, using the minimality of $\mathfrak{m}$, our assumption and Lemma 5.2 it will result that

$$\overline{\varprojlim \mathscr{A}_{I_\delta}} = \varprojlim \overline{\mathscr{A}_{I_\delta}},$$

for any $\delta < \alpha$, and

$$\overline{\varprojlim \mathscr{A}} = \overline{\varprojlim \mathscr{A}/P} = \varprojlim \overline{\mathscr{A}/P} = \varprojlim \overline{\mathscr{A}}/P = \varprojlim \overline{\mathscr{A}},$$

contradiction. ∎

References

[1] S. BREAZ and C. PELEA, Multialgebras and term functions over the algebra of their nonvoid subsets, *Mathematica (Cluj)* **43[66](2)** (2001), 143–149.

[2] P. CORSINI and V. LEOREANU, *Applications of hyperstructure theory*, Kluwer Academic Publishers, Boston–Dordrecht–London, 2003.

[3] G. GRÄTZER, A representation theorem for multi-algebras, *Arch. Math.* **3** (1962), 452–456.

[4] G. GRÄTZER, *Universal algebra*. Second edition, Springer-Verlag, 1979.

[5] G. HIGMAN and A. H. STONE, On inverse systems with trivial limits, *J. London Math. Soc.* **29** (1954), 233–236.

[6] V. LEOREANU, Direct limit and inverse limit of join spaces associated with fuzzy sets, *Pure Math. Appl.* **11(3)** (2000), 509–516.

[7] V. LEOREANU, Direct limits and inverse limits of SHR semigroups, *Southeast Asian Bulletin of Math.* **25** (2001), 421–426.

[8] B. MITCHELL, *Theory of categories*. Pure and Applied Mathematics, **Vol. XVII**, Academic Press, New York–London, 1965.

[9] C. PELEA, On the fundamental relation of a multialgebra, *Ital. J. Pure Appl. Math.* **10** (2001), 141–146.

[10] C. PELEA, Identities and multialgebras, *Ital. J. Pure Appl. Math.*, **15** (2004), 83–92.

[11] H. E. PICKETT, Homomorphisms and subalgebras of multialgebras, *Pacific J. Math.* **21** (1967), 327–342.

[12] I. PURDEA, *Tratat de algebră moderna.* **Vol. II.** (Romanian) [*Treatise on modern algebra.* **Vol. II**] With an English summary, Editura Academiei Republicii Socialiste România, Bucharest, 1982.

[13] D. SCHWEIGERT, Congruence relations of multialgebras, *Discrete Math.* **53** (1985), 249–253.

[14] T. VOUGIOUKLIS, On H_v-rings and H_v-representations, *Discrete Math.*, **208/209** (1999), 615–620.

[15] M. WALICKI and M. BIAŁASIK, Categories of relational structures, *Recent trends in algebraic development techniques* (Tarquinia, 1997), 418–433, *Lecture Notes in Comput. Sci.* **1376**, Springer, Berlin, 1998.

Cosmin Pelea

Babeş–Bolyai University
Faculty of Mathematics and Computer Science
Str. Mihail Kogălniceanu nr. 1
RO–3400 Cluj–Napoca, Romania
cpelea@math.ubbcluj.ro

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 65–89

ON M-CONTINUOUS FUNCTIONS AND PRODUCT SPACES

By

MARCELINA MOCANU

(Received August 2, 2004)

1. Introduction

Semi-open sets, preopen sets, α-sets and β-open sets in topological spaces play an important role in the researches on generalizations of continuity. By using these classes of sets, several authors introduced and studied various modifications of continuity, in the setting of topological spaces.

V. Popa and T. Noiri observed that the analogy among the results concerning these modifications of continuity suggests the need for a unified theory. They introduced and studied the fundamental notion of M-continuity [22], [23] and other related generalized forms of continuity: contra m-continuity [18], slightly m-continuity [24], weakly (τ, m)-continuity [25], weak M-continuity [26], faintly m-continuity [20]. The key concept of this approach is that of minimal structure, which led V. Popa and T. Noiri not only to a unified theory of the modifications of continuity, but also to new concepts and results. Another unified theory of several types of generalized continuity has been obtained by Császár [9], by using generalized topology.

In this paper, we investigate several types of minimal structures, especially minimal structures on a cartesian product of sets endowed with minimal structures. We obtain characterizations and properties of some types of M-continuous functions, unifying many known results concerning several classes of functions: D-continuous, D-supercontinuous, α-precontinuous, α-irresolute, semi-α-irresolute, α-preirresolute, β-preirresolute.

2000 Mathematics Subject Classification: 54C08

2. Preliminaries

In what follows, X and Y are always nonempty sets and $\mathcal{P}(X)$ (resp. $\mathcal{P}(Y)$) is the power set of X (resp. Y). If $\mathcal{C} \subseteq \mathcal{P}(X)$, $\mathcal{D} \subseteq \mathcal{P}(Y)$ and $f: X \to Y$, we will use the following notations: $f(\mathcal{C}) := \{f(A) : A \in \mathcal{C}\}$ and $f^{-1}(\mathcal{D}) := \{f^{-1}(B) : B \in \mathcal{D}\}$. We will denote by $\mathcal{U}(\mathcal{C})$ (resp. by $\mathcal{U}_F(\mathcal{C})$, $\mathcal{I}(\mathcal{C})$, $\mathcal{I}_F(\mathcal{C})$) the family of all unions (resp. finite unions, intersections, finite intersections) of all sets that belong to $\mathcal{C}$.

Let (X,τ) be a topological space and let A be a subset of X. The interior and the closure of A are denoted by $Int(A)$ and $Cl(A)$, respectively. The subset A is said to be semi-open [13] (resp. preopen [15], α-open [16], β-open [2]) if $A \subseteq Cl(Int(A))$ (resp. $A \subseteq Int(Cl(A))$, $A \subseteq Int(Cl(Int(A)))$, $A \subseteq Cl(Int(Cl(A)))$). The family of all semi-open (resp. preopen, α-open, β-open) sets in X is denoted by $SO(X)$ (resp. $PO(X)$, $\alpha(X)$, $\beta(X)$).

DEFINITION 2.1. (Popa and Noiri [23]) A subfamily m_X of $\mathcal{P}(X)$ is called a minimal structure (briefly m-structure) on X if $\emptyset \in m_X$ and $X \in m_X$. Each member of m_X is said to be m_X-open and the complement of an m_X-open set is said to be m_X-closed.

By (X, m_X) we denote a nonempty set X with a minimal structure m_X on X and we call (X, m_X) a space with minimal structure.

Let A be a subset of X. The m_X-interior of A, denoted by $m_X - Int(A)$, is the union of all m_X-open subsets of A, and the m_X-closure of A, denoted by $m_X - Cl(A)$, is the intersection of all m_X-closed supersets of A. The properties of operators of $m_X - Int$ and $m_X - Cl$ are stated in (Popa and Noiri [23], Lemma 3.1) and parallel the properties of operators Int and Cl of a topological space.

LEMMA 2.1. *(Popa and Noiri [25], Lemma 3.2) Let (X, m_X) be a space with minimal structure, let A be a subset of X and $x \in X$. Then $x \in m_X -$ $- Cl(A)$ if and only if $U \cap A \neq \emptyset$ for every $U \in m_X$ containing the point x.*

REMARK 2.1. Let m_X be a minimal structure on X. Then each of the families $\mathcal{U}(m_X)$, $\mathcal{I}(m_X)$, $\mathcal{U}_F(m_X)$, $\mathcal{I}_F(m_X)$, $\mathcal{U}(\mathcal{I}(m_X))$, $\mathcal{I}(\mathcal{U}(m_X))$, $\mathcal{U}_F(\mathcal{I}(m_X))$, $\mathcal{I}_F(\mathcal{U}(m_X))$, $\mathcal{U}_F(\mathcal{I}_F(m_X))$ and $\mathcal{I}_F(\mathcal{U}_F(m_X))$ is larger than m_X, in particular is a minimal structure on X. Denote by c_X the family of all m_X-closed sets. Then c_X is a minimal structure on X.

DEFINITION 2.2. (Popa and Noiri [23]) A minimal structure m_X on X is said to have property $(\mathcal{B})$ if every union of m_X-open sets is an m_X-open set.

We notice that m_X has property ($\mathscr{B}$) if and only if $\mathcal{U}(m_X) \subseteq m_X$.

LEMMA 2.2. *For every minimal structure m_X on X, we have $\mathcal{U}(m_X) = \{A : A \subseteq X,\ A \subseteq m_X - Int(A)\}$ and $\mathcal{I}(c_X) = \{B : B \subseteq X,\ m_X - Cl(B) \subseteq B\}$.*

PROOF. Denote $\mathcal{F} := \{A : A \subseteq X,\ A \subseteq m_X - Int(A)\}$. Then $\mathcal{F} = \{A : A \subseteq X,\ A = m_X - Int(A)\} \subseteq \mathcal{U}(m_X)$, since $m_X - Int(A) \subseteq A$ and $m_X - Int(A) = \bigcup\{U : U \subseteq A,\ U \in m_X\} \in \mathcal{U}(m_X)$, for every $A \subseteq X$. If $B \in \mathcal{U}(m_X)$, then there exist a subfamily $\{B_i : i \in I\}$ of m_X whose union is B. Then $\{B_i : i \in I\} \subseteq \{U : U \subseteq B,\ U \in m_X\}$, therefore applying the union it follows that $B \subseteq m_X - Int(B)$. Since $X \setminus m_X - Int(A) = m_X - Cl(X \setminus A)$ for every $A \subseteq X$, we have $\mathcal{I}(c_X) = \{X \setminus A : A \in \mathcal{U}(m_X)\} = \{X \setminus A : A \subseteq X,\ A \subseteq m_X - Int(A)\} = \{B : B \subseteq X,\ m_X - Cl(B) \subseteq B\}$. ∎

COROLLARY 2.1. *(Popa and Noiri [23], Lemma 3.2) For a minimal structure m_X on X the following are equivalent:*

(1) m_X has property ($\mathscr{B}$);

(2) If $m_X - Int(V) = V$, then $V \in m_X$;

(3) If $m_X - Cl(F) = F$, then $X \setminus F \in m_X$.

PROOF. By Lemma 2.2, (2) $\Leftrightarrow$ $\mathcal{U}(m_X) \subseteq m_X$ $\Leftrightarrow$ (1). The equivalence (2) $\Leftrightarrow$ (3) follows using the identities $m_X - Cl(X \setminus A) = X \setminus m_X - Int(A)$ and $m_X - Int(X \setminus A) = X \setminus m_X - Cl(A)$ for $A \subseteq X$ (Popa and Noiri [23], Lemma 3.1). ∎

The generalization of continuity in the setting of spaces with minimal structures is the fundamental notion of M-continuity, introduced by Popa and Noiri in [22], [23].

DEFINITION 2.3. A function $f\colon (X,m_X) \to (Y,m_Y)$ is said to be M-continuous at $x \in X$ if for each $V \in m_Y$ containing $f(x)$, there exists $U \in m_X$ containing x such that $f(U) \subseteq V$. A function $f\colon (X,m_X) \to (Y,m_Y)$ is said to be M-continuous if it is M-continuous at each point $x \in X$.

LEMMA 2.3. *(Popa and Noiri [23], Theorem 3.1) For a function $f\colon (X,m_X) \to (Y,m_Y)$ the following properties are equivalent:*

(1) *f is M-continuous;*

(2) *$f^{-1}(V) = m_X - Int(f^{-1}(V))$ for every $V \in m_Y$;*

(3) *$f(m_X - Cl(A)) \subseteq m_Y - Cl(f(A))$ for every subset A of X;*

(4) $m_X - Cl(f^{-1}(B)) \subseteq f^{-1}(m_Y - Cl(B))$ *for every subset* B *of* Y*;*

(5) $f^{-1}(m_Y - Int(B)) \subseteq m_X - Int(f^{-1}(B))$ *for every subset* B *of* Y*;*

(6) $m_X - Cl(f^{-1}(K)) = f^{-1}(K)$ *for every* m_Y*-closed set* K *of* Y*.*

LEMMA 2.4. *A function* $f:(X,m_X) \to (Y,m_Y)$ *is* M*-continuous if and only if* $f^{-1}(m_Y) \subseteq \mathcal{U}(m_X)$.

PROOF. We use the equivalence (1) $\Leftrightarrow$ (2) from Lemma 2.3 and the fact that, according to Lemma 2.2, (2) means that $f^{-1}(V) \in \mathcal{U}(m_X)$ for every $V \in m_Y$. ∎

COROLLARY 2.2. *A composition of* M*-continuous functions is* M*-continuous, if well-defined.*

PROOF. Let $f:(X,m_X) \to (Y,m_Y)$ and $g:(Y,m_Y) \to (Z,m_Z)$ be two M-continuous functions. Then the composition $g \circ f:(X,m_X) \to (Z,m_Z)$ is well-defined. Using Lemma 2.3 we obtain $(g \circ f)^{-1}(m_Z) = f^{-1}(g^{-1}(m_Z)) \subseteq$ $\subseteq f^{-1}(\mathcal{U}(m_Y)) = \mathcal{U}(f^{-1}(m_Y)) \subseteq \mathcal{U}(\mathcal{U}(m_X)) = \mathcal{U}(m_X)$. This shows that $g \circ f:(X,m_X) \to (Z,m_Z)$ is M-continuous, according to Lemma 2.4. ∎

PROPOSITION 2.1. *Let* X *and* Y *be nonempty sets with minimal structures* m_X *and* m_Y*, respectively, and let* $f:X \to Y$ *be a function. The following properties are equivalent:*

(1) $f:(X,m_X) \to (Y,m_Y)$ *is* M*-continuous;*

(2) $f:(X,m_X) \to (Y,\mathcal{U}(m_Y))$ *is* M*-continuous;*

(3) $f:(X,\mathcal{U}(m_X)) \to (Y,\mathcal{U}(m_Y))$ *is* M*-continuous.*

PROOF. We use Lemma 2.4. The implications (2) $\Rightarrow$ (1) and (3) $\Rightarrow$ (2) are obvious, since $m_Y \subseteq \mathcal{U}(m_Y)$ and $\mathcal{U}(\mathcal{U}(m_X)) = \mathcal{U}(m_X)$, respectively. Assume that (1) is true. Then $f^{-1}(\mathcal{U}(m_Y)) = \mathcal{U}(f^{-1}(m_Y)) \subseteq \mathcal{U}(\mathcal{U}(m_X))$, i.e. (3) is true. It follows that (1) $\Rightarrow$ (3). ∎

EXAMPLE 2.1. Let (X,τ) and (Y,σ) be topological spaces. Denote by $F(X)$, (resp. $F_\sigma(X)$, $G_\delta(X)$) the family of all closed (resp. F_σ, G_δ) subsets of X. Let $f:X \to Y$ be a function. If $f:(X,\tau) \to (Y,\sigma)$ is continuous, then it is easy to see that $f^{-1}(F(Y)) \subseteq F(X)$, $f^{-1}(\sigma \cap G_\delta(Y)) \subseteq \tau \cap$ $\cap\, G_\delta(X)$ and $f^{-1}(F(Y) \cap F_\sigma(Y)) \subseteq F(X) \cap F_\sigma(X)$, and so all the mappings $f:(X,F(X)) \to (Y,F(Y))$, $f:(X,\tau \cap G_\delta(X)) \to (Y,\widetilde{\tau} \cap G_\delta(Y))$ and $f:(X,F(X) \cap F_\sigma(X)) \to (Y,F(Y) \cap F_\sigma(Y))$ are M-continuous.

REMARK 2.2. V. Popa and T. Noiri have shown in [23] that semi-continuous, precontinuous, α-continuous, β-continuous, $s-\theta$-continuous, δ-almost continuous, δ-semi-continuous functions between topological spaces are examples of M-continuous functions.

Let (X,τ) and (Y,σ) be topological spaces. A function $f:(X,\tau)\to$ $\to (Y,\sigma)$ is said to be D-continuous if for each $x\in X$ and each open F_σ-set V of Y containing $f(x)$, there exists $U\in m_X$ containing x such that $f(U)\subseteq V$ (Kohli [11]).

Using Definition 2.3 we see that $f:(X,\tau)\to(Y,\sigma)$ is D-continuous if and only if $f:(X,\tau)\to(Y,\sigma\cap F_\sigma(Y))$ is M-continuous. It follows that Lemma 3.1 from (Kohli and Singh, [12]) is a consequence of Lemma 2.3.

$f:(X,\tau)\to(Y,\sigma)$ is H-almost continuous if and only if the inverse image of every open subset of Y is preopen in X [20]. We see that $f:(X,\tau)\to$ $\to(Y,\sigma)$ is H-almost continuous if and only if $f:(X,PO(X))\to(Y,\sigma)$ is M-continuous, by Lemma 2.4 and the fact that $PO(X)$ has property $(\mathscr{B})$.

REMARK 2.3. A function $f:(X,m_X)\to(Y,\sigma)$ is said to be contra m-continuous if $f^{-1}(V)=m_X-Cl(f^{-1}(V))$ for every open set V of Y, and $f:(X,m_X)\to(Y,\sigma)$ is said to be contra m-continuous at $x\in X$ if for each closed set F containing $f(x)$, there exists $U\in m_X$ containing x such that $f(U)\subseteq V$ (Noiri and Popa, [18]). We notice that $f:(X,m_X)\to(Y,\sigma)$ is contra m-continuous if and only if $f^{-1}(\sigma)\subseteq\mathcal{I}(c_X)$, according to Lemma 2.2. If $f:(X,m_X)\to(Y,\sigma)$ is contra m-continuous, then $f:(X,\mathcal{I}(c_X))\to(Y,\sigma)$ is M-continuous, by Lemma 2.4. The converse is not true, even if $m_X=\tau$ is a topology, since $\mathcal{I}(c_X)=\mathcal{I}(F(X))=F(X)$ is different of $\mathcal{U}(F(X))$ in general.

It follows, by Definition 2.3 and Lemma 2.3, that f is contra m-continuous at every point $x\in X$ if and only if $f:(X,m_X)\to(Y,F(Y))$ is M-continuous, i.e. $f^{-1}(F(Y))\subseteq\mathcal{U}(m_X)$, which is equivalent, by taking complementaries, to $f^{-1}(\sigma)\subseteq\mathcal{I}(c_X)$. Then f is contra m-continuous at every point $x\in X$ if and only if $f:(X,m_X)\to(Y,\sigma)$ is contra m-continuous (this is the equivalence (1) $\Leftrightarrow$ (3) in Theorem 3.2 from (Noiri and Popa, [18]).

Recall that for $A\subseteq Y$ the kernel of A is defined by $Ker(A)=\cap\{U\in$ $\in\sigma : A\subseteq U\}$. We notice that $Ker(A)=m_Y-Cl(A)$, where $m_Y=F(Y)$. Taking into account that $f:(X,m_X)\to(Y,\sigma)$ is contra m-continuous if and only if $f:(X,m_X)\to(Y,F(Y))$ is M-continuous, we see that Theorem 3.2 from (Noiri and Popa, [18]) follows from Lemma 2.3.

A function $f:(X,m_X) \to (Y,\sigma)$ is said to be faintly m-continuous [26] if for each $x \in X$ and each θ-open set V of Y containing $f(x)$, there is an open set U containing x such that $f(U) \subseteq V$ (Noiri and Popa [20]). The collection of θ-open sets in the topological space (Y,σ) forms a topology σ_θ on Y [27]. We notice that $f:(X,m_X) \to (Y,\sigma)$ is faintly m-continuous if and only if $f:(X,m_X) \to (Y,\sigma_\theta)$ is M-continuous.

DEFINITION 2.4. A function $f:(X,m_X) \to (Y,m_Y)$ is said to be M-open if $f(U) \in m_Y$ for every $U \in m_X$, respectively is said to be almost M-open if $f(m_X - Int(A)) \subseteq m_Y - Int(f(A))$ for every subset A of X.

REMARK 2.4. Let (X,τ_X) and (Y,τ_Y) be topological spaces. If $m_Y = \tau_Y$ (resp. $m_Y = SO(X)$, $m_Y = \alpha(Y)$, $m_Y = \beta(Y)$), then an M-open function $f:(X,\tau_X) \to (Y,m_Y)$ is called open (resp. semi-open [17], α-open [15], β-open [2]).

LEMMA 2.5. *A function $f:(X,m_X) \to (Y,m_Y)$ is M-open (resp. almost M-open if and only if $f(m_X) \subseteq m_Y$ (resp. $f(m_X) \subseteq \mathcal{U}(m_Y)$.*

PROOF. The characterization of M-open functions is clear from Definition 2.4. Let $f:(X,m_X) \to (Y,m_Y)$ be almost M-open and let $U \in m_X$. Then $U = m_X - Int(U)$, hence $f(U) = f(m_X - Int(U)) \subseteq m_Y - Int(f(U))$. It follows by Lemma 2.2 that $f(U) \in \mathcal{U}(m_Y)$, hence $f(m_X) \subseteq \mathcal{U}(m_Y)$. For the converse, assume that $f(m_X) \subseteq \mathcal{U}(m_Y)$. Let A be an arbitrary subset of X. Denote $B = f(m_X - Int(A))$. Then $B = f(\bigcup\{U : U \in m_X, U \subseteq \subseteq A\}) = \bigcup\{f(U) : U \in m_X, U \subseteq A\}$. By our assumption, $f(U) \in \mathcal{U}(m_Y)$ for each $U \in m_X$. This yields $B \in \mathcal{U}(\mathcal{U}(m_Y)) = \mathcal{U}(m_Y)$. By Lemma 2.2, $B \subseteq m_Y - Int(B)$. ∎

COROLLARY 2.3. *Every M-open function $f:(X,m_X) \to (Y,m_Y)$ is almost M-open.*

COROLLARY 2.4. *Let $f:(X,m_X) \to (Y,m_Y)$ be a bijective function. Then f is almost M-open if and only if f^{-1} is M-continuous.*

PROOF. By Lemma 2.4, $f^{-1} : (Y,m_Y) \to (X,m_X)$ is M-continuous if and only if $(f^{-1})^{-1}(m_X) \subseteq \mathcal{U}(m_Y)$, i.e. $f(m_X) \subseteq \mathcal{U}(m_Y)$, which is equivalent to the property of $f:(X,m_X) \to (Y,m_Y)$ to be almost M-open, according to Lemma 2.5. ∎

3. Some types of minimal structures

In this section we study some useful examples of minimal structures.

Given a topological space (X,τ), let T be any finite composition of operators *Int* and *Cl*. Denote $GO_T(X) = \{S \subseteq X : S \subseteq T(S)\}$.

LEMMA 3.1. *If (X,τ) is a topological space and T is a finite composition of operators Int and Cl, then $GO_T(X)$ is a minimal structure which has property $(\mathscr{B})$ and $\tau \subseteq GO_T(X)$.*

PROOF. Let $T = T_1 \circ T_2 \circ \cdots \circ T_n$, where $T_k \in \{Int, Cl\}$, $k \in \{1,2,\ldots,n\}$. Since for every $A \subseteq X$ we have $Int(A) \subseteq T_k(A)$ for $k \in \{1,2,\ldots \ldots,n\}$, and $S \subseteq Int(S)$ for every $S \in \tau$, it follows that $\tau \subseteq GO_T(X)$, hence $GO_T(X)$ is a minimal structure. Let $\{A_i : i \in I\} \subseteq \mathscr{P}(X)$. Obviously, $\bigcup\{Int(A_i) : i \in I\} \subseteq Int\left(\bigcup\{A_i : i \in I\}\right)$ and $\bigcup\{Cl(A_i) : i \in I\} \subseteq \subseteq Cl\left(\bigcup\{A_i : i \in I\}\right)$, whence it follows by induction on n that $\bigcup\{T(A_i) : i \in I\} \subseteq T\left(\bigcup\{A_i : i \in I\}\right)$. Then, for every family $\{S_i : i \in I\} \subseteq GO_T(X)$ we have $\bigcup\{S_i : i \in I\} \subseteq \bigcup\{T(S_i) : i \in I\} \subseteq T\left(\bigcup\{S_i : i \in I\}\right)$, hence $\bigcup\{S_i : i \in I\} \in GO_T(X)$. This shows that $GO_T(X)$ has property $(\mathscr{B})$. ∎

REMARK 3.1. If for each $A \subseteq X$ we denote $T(A) = Int(A)$ (resp. $T(A) = = Cl(A)$, $T(A) = Cl(Int(A))$, $T(A) = Int(Cl(A))$, $T(A) = Int(Cl(Int(A)))$, $T(A) = Cl(Int(Cl(A))))$, then $GO_T(X) = \tau$ (resp. $GO_T(X) = \mathscr{P}(X)$, $GO_T(X) = SO(X)$, $GO_T(X) = PO(X)$, $GO_T(X) = \alpha(X)$, $GO_T(X) = = \beta(X)$). By Lemma 3.1, the families $SO(X)$, $PO(X)$, $\alpha(X)$, $\beta(X)$, are all minimal structures with property $(\mathscr{B})$.

THEOREM 3.1. *Let (X,τ_X) and (Y,τ_Y) be topological spaces and let $f: X \to Y$. Let T be any finite composition of operators Int and Cl. If $f:(X,\tau_X) \to (Y,\tau_Y)$ is continuous and open, then $f:(X, GO_T(X)) \to \to (Y, GO_T(Y))$ is M-continuous and M-open.*

PROOF. Let $A \subseteq X$. The continuity of $f:(X,\tau_X) \to (Y,\tau_Y)$ implies $f(Cl(A)) \subseteq Cl(f(A))$. Since $f:(X,\tau_X) \to (Y,\tau_Y)$ is open, we have $f(Int(A)) \subseteq Int(f(A))$. Then $f(T(A)) \subseteq T(f(A))$. If $A \in GO_T(X)$, then $f(A) \subseteq f(T(A)) \subseteq T(f(A))$, that is, $f(A) \in GO_T(Y)$. This proves that $f:(X, GO_T(X)) \to (Y, GO_T(Y))$ is M-open.

Let $B \subseteq Y$. The continuity of $f:(X,\tau_X) \to (Y,\tau_Y)$ implies $f^{-1}(Int(B)) \subseteq \subseteq Int(f^{-1}(B))$ and $Cl(f^{-1}(B)) \subseteq f^{-1}(Cl(B))$. Since $f:(X,\tau_X) \to (Y,\tau_Y)$ is open, it follows that:

i) $f(Int(f^{-1}(B))) \subseteq Int(f(f^{-1}(B))) \subseteq Int(B)$ and consequently $f^{-1}(Int(B)) \supseteq Int(f^{-1}(B))$;

ii) $X \setminus f^{-1}(Cl(B)) = f^{-1}(Y \setminus Cl(B)) = f^{-1}(Int(B)) \supseteq Int(f^{-1}(B))$, hence $f^{-1}(Cl(B)) \subseteq X \setminus Int(f^{-1}(B)) = Cl(f^{-1}(B))$.

We conclude that the operator f^{-1} commutes with the operators *Int* and *Cl*: $f^{-1}(Int(B)) = Int(f^{-1}(B))$ and $f^{-1}(Cl(B)) = Cl(f^{-1}(B))$ for every $B \subseteq Y$. It follows that $f^{-1}(T(B)) = T(f^{-1}(B))$ for every $B \subseteq Y$. If $B \in GO_T(Y)$ then $f^{-1}(B) \subseteq f^{-1}(T(B)) = T(f^{-1}(B))$, hence $f^{-1}(B) \in \in GO_T(X)$. This proves that $f:(X, GO_T(X)) \to (Y, GO_T(Y))$ is M-continuous, according to Lemma 2.4. ∎

Let X be a nonempty set with a minimal structure m_X. Since the union of all members of m_X is X, we notice that m_X is a subbase for a topology on X. There exists a unique smallest topology including m_X, called the topology generated by m_X, which we denote by $\tau(m_X)$. Namely, $\tau(m_X) = \mathcal{U}(\mathcal{I}_F(m_X))$ and m_X is a subbase for the topology $\tau(m_X)$. In many important cases, a minimal structure is not closed to intersection.

EXAMPLE 3.1. Recall that in every topological space (X, τ) we have $SO(X) \cup PO(X) \subseteq \beta(X)$.

Let $X = \{a, b, c, d\}$ and $\tau = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, b, c\}, X\}$. Then τ is a topology on X. Let $S_1 = \{a, c, d\}$ and $S_2 = \{b, d\}$. Then S_1 and S_2 are semi-open sets with respect to τ, since $Cl(Int(S_k)) = S_k$, $k = 1, 2$, while $S_1 \cap S_2 = \{d\}$ is not β-open, because $Cl(Int(Cl(S_1 \cap S_2))) = \emptyset$. This example shows that $SO(X)$ and $\beta(X)$ are not closed to intersection.

Now we study minimal structures having a weaker property than that to be closed to finite intersections.

DEFINITION 3.1. A minimal structure m_X on X is said to have property $(\mathcal{I}_1)$ if for every $U, V \in m_X$ such that $U \cap V \neq \emptyset$ and for each $x \in U \cap V$, there exists $W \in m_X$ such that $x \in W \subseteq U \cap V$.

REMARK 3.2. If m_X has property $(\mathcal{I}_1)$, $n \geq 1$ and $U_k, k \in \{1, 2, \ldots \ldots, n\}$, are m_X-open sets having nonempty intersection, then for each $x \in \in \cap\{U_k : k \in \{1, 2, \ldots, n\}\}$ there exists $W \in m_X$ such that $x \in W \subseteq \subseteq \cap\{U_k : k = \overline{1, n}\}$.

LEMMA 3.2. *For a minimal structure m_X on X the following properties are equivalent:*

(1) m_X has property $(\mathcal{I}_1)$;

(2) $\mathcal{I}_F(m_X) \subseteq \mathcal{U}(m_X)$;

(3) $\mathcal{I}_F(\mathcal{U}(m_X)) \subseteq \mathcal{U}(m_X)$;

(4) $\mathcal{U}(m_X)$ is a topology on X (the smallest topology on X including m_X);

(5) m_X is a base for a topology.

PROOF. (1) $\Rightarrow$ (2): Let $U, V \in m_X$. Assume that $U \cap V \neq \emptyset$ (otherwise, $U \cap V = \emptyset \in m_X \subseteq \mathcal{U}(m_X)$). By (1), for each $x \in U \cap V$ there exists $W_x \in m_X$ such that $x \in W_x \subseteq U \cap V$. Then $U \cap V = \bigcup\{W_x : x \in U \cap \cap V\} \in \mathcal{U}(m_X)$. If m_X has property $(\mathcal{I}_1)$, we proved that $U, V \in m_X$ implies $U \cap V \in \mathcal{U}(m_X)$. Now (2) follows by induction.

(2) $\Rightarrow$ (3): We have $\mathcal{I}_F(\mathcal{U}(\mathcal{C})) \subseteq \mathcal{U}(\mathcal{I}_F(\mathcal{C}))$ for every $\mathcal{C} \subseteq \mathcal{P}(X)$. Then (2) implies $\mathcal{I}_F(\mathcal{U}(m_X)) \subseteq \mathcal{U}(\mathcal{I}_F(m_X)) \subseteq \mathcal{U}(\mathcal{U}(m_X)) = \mathcal{U}(m_X)$.

(3) $\Rightarrow$ (4): $\{\emptyset, X\} \subseteq m_X \subseteq \mathcal{U}(m_X)$. Obviously, $\mathcal{U}(m_X)$ is closed to arbitrary unions (has property $(\mathcal{B})$). By (3), $\mathcal{U}(m_X)$ is closed to finite intersections.

(4) $\Rightarrow$ (5): By the definition of a base of a given topology, m_X is a base for the topology $\mathcal{U}(m_X)$.

(5) $\Rightarrow$ (1): Let τ be a topology on X having m_X as a base, i.e. $m_X \subseteq \subseteq \tau \subseteq \mathcal{U}(m_X)$. Suppose that $U, V \in m_X$ and $x \in U \cap V$. Then $U \cap V \in \tau$, therefore $U \cap V$ is a union of m_X-open sets and at least one of these sets contains the point x. ∎

LEMMA 3.3. *Let X be a nonempty set with a minimal structure m_X. Then the following statements are equivalent:*

(1) m_X is a base for a topology on X;

(2) If $A \subseteq X$ and $U, V \in m_X$ satisfy the condition $(U \cap V) \cap m_X - - Cl(A) \neq \emptyset$, then $U \cap V \cap A \neq \emptyset$.

PROOF. (1) $\Rightarrow$ (2): Let $A \subseteq X$ and $U, V \in m_X$ satisfying the condition $(U \cap V) \cap m_X - Cl(A) \neq \emptyset$. Pick $x \in (U \cap V) \cap m_X - Cl(A) \neq \emptyset$. According to Lemma 3.2, m_X has property $(\mathcal{I}_1)$, hence there is $W \in m_X$ such that $x \in W \subseteq U \cap V$. Since $x \in m_X - Cl(A)$, we have $W \cap A \neq \emptyset$, by Lemma 2.1, hence $U \cap V \cap A \neq \emptyset$.

(2) $\Rightarrow$ (1): We assume that (1) is false and we prove that (2) is false. For each $x \in X$, denote $\mathcal{V}(x) = \{V : V \in m_X, x \in V\}$. We notice that $\mathcal{V}(x)$ is nonempty, since $X \in \mathcal{V}(x)$.

By our assumption, there exist two m_X-open sets U_0, V_0 and $x_0 \in$ $\in U_0 \cap V_0$ such that no member of $\mathcal{V}(x_0)$ is included in $U_0 \cap V_0$. For each $W \in \mathcal{V}(x_0)$ we find a point $x_W \in W \setminus (U_0 \cap V_0)$. Put $A = \{x_W : W \in \mathcal{V}(x_0)\}$. Then $A \subseteq X \setminus (U_0 \cap V_0)$. For every $W \in \mathcal{V}(x_0)$ we have $x_W \in W \cap A$, hence $x_0 \in m_X - Cl(A)$. We have $x_0 \in (U_0 \cap V_0) \cap m_X - Cl(A)$, but $U_0 \cap V_0 \cap A = \emptyset$. This shows that (2) is false. ∎

EXAMPLE 3.2. *Let* $X = \{a, b, c\}$ *and* $m_X = \{\emptyset, X, \{a, b\}, \{b, c\}\}$. *Then* $\mathcal{I}_F(m_X) = m_X \cup \{\{b\}\}$ *and* $\mathcal{U}(m_X) = m_X$. *It follows by Lemma 3.2 that* m_X *does not have property* $(\mathcal{I}_1)$. *Let* $A = \{a, c\}$. *Then* $b \in m_X - Cl(A)$. *Let* $U = \{a, b\}$, $V = \{b, c\}$. *Then* $b \in (U \cap V) \cap m_X - Cl(A)$, *but* $U \cap V \cap A = \emptyset$.

Next we attach to any given minimal structure m_X on a nonempty set X another minimal structure $Inc(m_X)$, defined as follows: we say that $A \in Inc(m_X)$ if for every $x \in A$ there exists $B \in m_X$ such that $x \in B \subseteq A$.

PROPOSITION 3.1. $Inc(m_X) = \mathcal{U}(m_X)$.

PROOF. $Inc(m_X) \subseteq \mathcal{U}(m_X)$: Let $A \in Inc(m_X)$. For every $x \in A$ there exists $B_x \in m_X$ such that $x \in B_x \subseteq A$. Then $A = \bigcup\{\{x\} : x \in A\} \subseteq$ $\subseteq \bigcup\{B_x : x \in A\} \subseteq A$, hence $A = \bigcup\{B_x : x \in A\} \in \mathcal{U}(m_X)$.

$\mathcal{U}(m_X) \subseteq Inc(m_X)$: If $A \in \mathcal{U}(m_X)$, then $A = \bigcup\{A_i : i \in I\}$ for some family $\{A_i : i \in I\}$ of m_X-open sets. For every $x \in A$ there is $i(x) \in I$ such that $x \in A_{i(x)}$. This shows that $A \in Inc(m_X)$. ∎

COROLLARY 3.1. *For every minimal structure* m_X *on* X, *the minimal structure* $Inc(m_X)$ *has property* $(\mathcal{B})$.

REMARK 3.3. Let (Y, σ) be a topological space. Denote by $m_Y = CO(Y)$ the family of all clopen subsets of Y (sets which are simultaneously open and closed in Y). A subset G of Y is said to be δ^*-open if for each $y \in G$ there exists a clopen set V such that $y \in V \subseteq G$. According to Proposition 3.1, the family of all δ^*-open subsets of Y is $Inc(CO(Y)) = \mathcal{U}(CO(Y))$. Since $\mathcal{I}_F(CO(Y)) = CO(Y)$, the set of all δ^*-open subsets of Y is a topology, called the ultra-regularization of σ, and denoted by σ_u [24]. A function $f : (X, m_X) \to (Y, \sigma)$ is said to be slightly m-continuous if $f : (X, m_X) \to$ $\to (Y, CO(Y))$ is M-continuous (Popa and Noiri [24]). By Applying Proposition 2.1, Lemma 2.3 and Lemma 2.4 it follows that $f : (X, m_X) \to (Y, \sigma)$ is slightly m-continuous if and only if $f^{-1}(G) = m_X - Int(f^{-1}(G))$ for every δ^*-open set $G \subseteq Y$ (which is obviously equivalent to the fact that

$f^{-1}(K) = m_X - Cl(f^{-1}(K))$ for every δ^*-closed set $K \subseteq Y$). This result is Theorem 3.1 from (Popa and Noiri [24]).

REMARK 3.4. A set G in a topological space (X,τ) is said to be d-open if for each $x \in G$, there exists an open F_σ-set H such that $x \in H \subseteq G$ (Kohli and Singh [12]). It follows, by Proposition 3.1, that the family of d-open sets is $Inc(\tau \cap F_\sigma(X)) = \mathcal{U}(\tau \cap F_\sigma(X)) \subseteq \tau$. A function $f\colon X \to Y$ from a topological space (X,τ) into a topological space $(Y,\widetilde{\tau})$ is D-supercontinuous if and only if inverse image of every open subset of Y is d-open in X (Kohli and Singh, Theorem 3.1). Obviously, every D-supercontinuous function is continuous. Lemma 2.4 shows that $f\colon (X,\tau) \to (Y,\widetilde{\tau})$ is D-supercontinuous if and only if $f\colon (X,\tau \cap F_\sigma(X)) \to (Y,\widetilde{\tau})$ is M-continuous. We notice that Theorems 3.1, 3.2 and 3.3 from (Kohli and Singh [12]) are consequences of Lemma 2.3, and Theorems 3.6 and 3.10 from (Kohli and Singh [12]) follow from Corollary 2.2.

We can introduce the notion of *quotient minimal structure*, a generalization of the quotient topology.

Let $f\colon X \to Y$ be a function from a topological space (X,τ_X) onto a set Y. The family $\tau_Y \subseteq \mathcal{P}(Y)$, defined by $V \in \tau_Y$ if and only if $f^{-1}(V) \in \tau_X$, is a topology on Y, called the quotient topology (induced by f). It turns out that the quotient topology above is the finest topology τ_Y for which $f\colon (X,\tau_X) \to (Y,\tau_Y)$ is continuous.

In the case where $f\colon X \to Y$ is a function from a space with minimal structure (X,m_X) onto a set Y, we look for the largest minimal structure m_Y on Y for which $f\colon (X,m_X) \to (Y,m_Y)$ is M-continuous.

DEFINITION 3.2. Let $f\colon X \to Y$ be a function from a space with minimal structure (X,m_X) onto a set Y. The minimal structure $m_Y^Q \subseteq \mathcal{P}(Y)$, defined by $V \in m_Y^Q$ if and only if $f^{-1}(V) \in \mathcal{U}(m_X)$ is called the quotient minimal structure (induced by f) on Y.

REMARK 3.5. Obviously, the family m_Y^Q defined as above is indeed a minimal structure, which has property ($\mathcal{B}$). Furthermore, $m_Y^Q = f(f^{-1}(m_Y^Q)) \subseteq$ $\subseteq f(\mathcal{U}(m_X)) = \mathcal{U}(f(m_X))$. If m_X is a topology, then m_Y^Q is a topology.

THEOREM 3.2. *The quotient minimal structure induced by a surjective function $f\colon (X,m_X) \to Y$ is the largest minimal structure m_Y on Y for which $f\colon (X,m_X) \to (Y,m_Y)$ is M-continuous.*

PROOF. The M-continuity of $f:(X,m_X)\to(Y,m_Y^Q)$ follows by Definition 3.2 and Lemma 2.4. Let m_Y be a minimal structure on Y, such that $f:(X,m_X)\to(Y,m_Y)$ is M-continuous. For every $V\in m_Y$, we have $f^{-1}(V)\in\mathcal{U}(m_X)$, hence $V\in m_Y^Q$, by Definition 3.2. ∎

THEOREM 3.3. *Let $f:X\to Y$ be a function from a space with minimal structure (X,m_X) onto a set Y and let m_Y^Q be the quotient minimal structure induced by f. Then $g:(Y,m_Y^Q)\to(Z,m_Z)$ is M-continuous if and only if $g\circ f:(X,m_X)\to(Z,m_Z)$ is M-continuous.*

PROOF. By Theorem 3.2, $f:(X,m_X)\to(Y,m_Y^Q)$ is M-continuous. The necessity follows by Corollary 2.2. Assuming that $g\circ f:(X,m_X)\to(Z,m_Z)$ is M-continuous, for each $W\in m_Z$ we have $(g\circ f)^{-1}(W)=f^{-1}(g^{-1}(W))\in$ $\in\mathcal{U}(m_X)$, hence $g^{-1}(W)\in m_Y^Q$. ∎

REMARK 3.6. Let (X,τ) be a topological space, $m_X=\tau\cap F_\sigma(X)$ and let $f:X\to Y$. The quotient minimal structure induced by f on Y is the D-quotient topology introduced by Kohli and Singh [12]. In this setting, Theorem 3.2 and Theorem 3.3 give, respectively, Theorem 4.1 and Theorem 4.2 from [12].

4. Minimal structures on a cartesian product

Let $\{(X_\alpha,m_\alpha):\alpha\in\Delta\}$ be a family of spaces with minimal structure. Let $X:=\prod\{X_\alpha:\alpha\in\Delta\}$ be the cartesian product of the sets X_α, $\alpha\in\Delta$. For each $\alpha\in\Delta$ denote by $\pi_\alpha:X\to X_\alpha$ the canonical projection of X onto X_α.

We consider the families of subsets of X defined by $\mathcal{S}_\alpha:=\{\pi_\alpha^{-1}(U_\alpha):U_\alpha\in m_\alpha\}$, $\alpha\in\Delta$, and $\mathcal{S}=\bigcup\{\mathcal{S}_\alpha:\alpha\in\Delta\}$. Notice that $\mathcal{S}$ is always a subbase for a topology, since the union of all members of $\mathcal{S}$ is X.

If $m_\alpha=\tau_\alpha$ is a topology on X_α for each $\alpha\in\Delta$, then the product topology on X is the smallest topology τ on X for which all projections $\pi_\alpha:(X,\tau)\to(X_\alpha,\tau_\alpha)$, $\alpha\in\Delta$, are continuous. It turns out that the product topology τ on X is the topology generated by $\mathcal{S}$, i.e. $\tau=\mathcal{U}(\mathcal{I}_F(\mathcal{S}))$.

In the general case, it would be interesting to find the smallest minimal structure m_X on X for which all projections $\pi_\alpha:(X,m_X)\to(X_\alpha,m_\alpha)$, $\alpha\in\Delta$, are M-continuous; this condition is satisfied if and only if $\mathcal{S}\subseteq\mathcal{U}(m_X)$. Denote by $\mathcal{M}$ the collection of all minimal structures m_X on X satisfying the condition $\mathcal{S} \subseteq \mathcal{U}(m_X)$. Unfortunately, a smallest element of $(\mathcal{M}, \subseteq)$ need not exist in general, as the following example shows.

EXAMPLE 4.1. Let $A = \{a, b\}$. Let us take $X_k = A$ with the minimal structure $m_k = \mathcal{P}(A)$, for $k = 1, 2$. Then

$$\mathcal{S} = \{\emptyset, A \times A, \{a\} \times A, \{b\} \times A, A \times \{a\}, A \times \{b\}\}.$$

We consider the following minimal structures on $X = X_1 \times X_2$: $m_X^1 = \mathcal{S}$ and $m_X^2 = \{\emptyset, A \times A, \{(a,a)\}, \{(a,b)\}, \{(b,a)\}, \{(b,b)\}\}$. Then $\mathcal{S} \subseteq \mathcal{U}(m_X^k)$, hence $m_X^k \in \mathcal{M}$, for $k = 1, 2$. If $(\mathcal{M}, \subseteq)$ would have a smallest element m_X, then $m_X \subseteq m_X^1 \cap m_X^2 = \{\emptyset, A \times A\}$, hence $\mathcal{U}(m_X) \subseteq \{\emptyset, A \times A\}$. We notice that $\mathcal{S} \nsubseteq \mathcal{U}(m_X)$, a contradiction with the assumption $m_X \in \mathcal{M}$.

In what follows, we will endow X with the topology generated by the subbase $\mathcal{S}$, which is the smallest (coarsest) topology on X for which all projections are M-continuous.

DEFINITION 4.1. The topology determined by a family $\{(X_\alpha, m_\alpha) : \alpha \in \Delta\}$ of spaces with minimal structure, on the cartesian product $X := \prod\{X_\alpha : \alpha \in \Delta\}$, is the smallest topology τ_X for which all projections $\pi_\alpha : (X, \tau_X) \to (X_\alpha, m_\alpha)$, $\alpha \in \Delta$, are M-continuous.

REMARK 4.1. a) The topology τ_X in the definition above is given by $\tau_X = \mathcal{U}(\mathcal{I}_F(\mathcal{S}))$.

b) We also can say that the projections $\pi_\alpha : (X, \tau_X) \to (X_\alpha, m_\alpha)$, $\alpha \in \Delta$, are (τ, m)-continuous. (Popa and Noiri [25]).

DEFINITION 4.2. (Noiri and Popa [19]) A nonempty set X with a minimal structure m_X, (X, m_X), is said to be m-Hausdorff if for each distinct points $x, y \in X$, there exist $U, V \in m_X$ containining x and y, respectively, such that $U \cap V = \emptyset$.

The following two properties are natural generalizations of well-know results on product topological spaces.

THEOREM 4.1. *If $\mathcal{F} = \{(X_\alpha, m_\alpha) : \alpha \in \Delta\}$ is a family of m-Hausdorff spaces and τ_X is the topology determined by $\mathcal{F}$ on $X := \prod\{X_\alpha : \alpha \in \Delta\}$, then the topological space (X, τ_X) is Hausdorff.*

PROOF. Let x and y be distinct points in X. There exists $\alpha \in \Delta$ such that $\pi_\alpha(x) \neq \pi_\alpha(y)$. Since (X_α, m_α) is m-Hausdorff, there exist two disjoint m_α-open sets U_α, V_α, containing $\pi_\alpha(x)$ and $\pi_\alpha(y)$, respectively. Then $\pi_\alpha^{-1}(U_\alpha)$

and $\pi_\alpha^{-1}(V_\alpha)$ are disjoint sets in the family $\mathcal{S}_\alpha \subseteq \mathcal{S} \subseteq \tau_X$, containing x and y respectively. ∎

THEOREM 4.2. *Let $\mathcal{F} = \{(X_\alpha, m_\alpha) : \alpha \in \Delta\}$ be a family of spaces with minimal structure and let τ_X be the topology determined by $\mathcal{F}$ on $X := \prod\{X_\alpha : \alpha \in \Delta\}$. Then, for every family of sets $\{A_\alpha : \alpha \in \Delta, A_\alpha \subseteq \subseteq X_\alpha\}$ we have*

$$\tau_X\text{-}Cl\Big(\prod_{\alpha\in\Delta} A_\alpha\Big) \subseteq \prod_{\alpha\in\Delta} m_\alpha\text{-}Cl(A_\alpha).$$

PROOF. Let $x \in \tau_X - Cl\Big(\prod_{\alpha\in\Delta} A_\alpha\Big)$. Pick an arbitrary $\beta \in \Delta$. Let us take $U_\beta \in m_\beta$ containing $x_\beta = \pi_\beta(x)$. Since $\pi_\beta^{-1}(U_\beta)$ is in τ_X and contains the point x, it follows by Lemma 2.1 that $\pi_\beta^{-1}(U_\beta) \cap \prod_{\alpha\in\Delta} A_\alpha \neq \emptyset$, which is equivalent to $U_\beta \cap A_\beta \neq \emptyset$. We proved that each set of m_β containing x_β meets A_β, hence $x_\beta \in m_\beta - Cl(A_\beta)$, by Lemma 2.1. ∎

If $m_\alpha = \tau_\alpha$ is a topology on X_α for each $\alpha \in \Delta$, it is known that the inclusion $\tau_X - Cl\Big(\prod_{\alpha\in\Delta} A_\alpha\Big) \supseteq \prod_{\alpha\in\Delta} m_\alpha - Cl(A_\alpha)$ also holds, for every family of sets $\{A_\alpha : \alpha \in \Delta, A_\alpha \subseteq X_\alpha\}$. We will establish a necessary and sufficient condition for the inclusion above, in the general case where $\{(X_\alpha, m_\alpha) : \alpha \in \Delta\}$ is a family of spaces with minimal structure.

THEOREM 4.3. *Let $\mathcal{F} = \{(X_\alpha, m_\alpha) : \alpha \in \Delta\}$ be a family of spaces with m-structure and let τ_X be the topology determined by $\mathcal{F}$ on $X := \prod\{X_\alpha : \alpha \in \in \Delta\}$. The following properties are equivalent:*

(1) $\tau_X - Cl\Big(\prod_{\alpha\in\Delta} A_\alpha\Big) \supseteq \prod_{\alpha\in\Delta} m_\alpha - Cl(A_\alpha)$ for every family of sets $\{A_\alpha : \alpha \in \Delta, A_\alpha \subseteq X_\alpha\}$;

(2) For each $\alpha \in \Delta$ the minimal structure m_α is a base for a topology on X_α.

PROOF. (1) $\Rightarrow$ (2): Fix $\beta \in \Delta$ and let A_β be an arbitrary subset of X_β. Take U_β, V_β be m_β-open sets such that $(U_\beta \cap V_\beta) \cap m_\beta - Cl(A_\beta) \neq \emptyset$. We shall prove that $U_\beta \cap V_\beta \cap A_\beta \neq \emptyset$, which shows, by Lemma 3.3, that m_β is a base for a topology on X_β. Put $A_\alpha = X_\alpha$ for every $\alpha \in \Delta\backslash\{\beta\}$. Pick $a_\beta \in (U_\beta \cap V_\beta) \cap$ $\cap\, m_\beta - Cl(A_\beta)$ and $a_\alpha \in X_\alpha$ for each $\alpha \in \Delta \setminus \{\beta\}$. Let $a = (a_\alpha)_{\alpha\in\Delta}$. Then $a \in \prod_{\alpha\in\Delta} m_\alpha - Cl(A_\alpha)$, hence by (1) we have $a \in \tau_X - Cl\Big(\prod_{\alpha\in\Delta} A_\alpha\Big)$. Let $W := (U_\beta \cap V_\beta) \times \prod_{\alpha\neq\beta} X_\alpha$. Then $W = \pi_\beta^{-1}(U_\beta) \cap \pi_\beta^{-1}(V_\beta) \in \mathcal{I}_F(\mathcal{S}) \subseteq \tau_X$ and W contains a. By Lemma 2.1, we have $W \cap \prod_{\alpha\in\Delta} A_\alpha \neq \emptyset$, which is equivalent to $U_\beta \cap V_\beta \cap A_\beta \neq \emptyset$.

(2) $\Rightarrow$ (1): Let $\{A_\alpha : \alpha \in \Delta,\ A_\alpha \subseteq X_\alpha\}$ be a family of sets such that A_α is nonempty for each $\alpha \in \Delta$ (if A_α is empty for some $\alpha \in \Delta$, then the claim is obvious). Take $x \in \prod_{\alpha\in\Delta} m_\alpha - Cl(A_\alpha)$, i.e. $x_\alpha := \pi_\alpha(x) \in m_\alpha - Cl(A_\alpha)$ for each $\alpha \in \Delta$. Let $V \in \tau_X$ containing x. Since $\tau_X = \mathcal{U}(\mathcal{I}_F(\mathcal{S}))$, there exists $U \in \mathcal{I}_F(\mathcal{S})$ such that $x \in U \subseteq V$. We can write U under the form $U =$ $= \prod_{k=1}^{n} U_{\alpha_k} \times \prod_{\alpha\in\Phi} X_\alpha$, where $n \geq 1$, $\{\alpha_1,\alpha_2,\ldots,\alpha_n\} \subseteq \Delta$, $\Phi = \Delta \setminus \{\alpha_1,\alpha_2,\ldots$ $\ldots,\alpha_n\}$ and $U_{\alpha_k} \in \mathcal{I}_F(m_{\alpha_k})$, $k = \overline{1,n}$. We apply the fact that (2) implies, by Lemma 3.2, that m_α has property $(\mathcal{I}_1)$, for every $\alpha \in \Delta$. Then for each $k \in \{1,2,\ldots,n\}$ there exists $W_{\alpha_k} \in m_{\alpha_k}$ such that $x_{\alpha_k} \in W_{\alpha_k} \subseteq U_{\alpha_k}$. Since $x_{\alpha_k} \in m_{\alpha_k} - Cl(A_{\alpha_k})$, by Lemma 2.1 we have $W_{\alpha_k} \cap A_{\alpha_k} \neq \emptyset$, for each $k \in \{1,2,\ldots,n\}$. Define the set $P = \prod_{k=1}^{n} (W_{\alpha_k} \cap A_{\alpha_k}) \times \prod_{\alpha\in\Phi} X_\alpha$. Then

$$\emptyset \neq P = \Big(\prod_{k=1}^{n} W_{\alpha_k} \times \prod_{\alpha\in\Phi} X_\alpha\Big) \cap \prod_{\alpha\in\Delta} A_\alpha \subseteq U \cap \prod_{\alpha\in\Delta} A_\alpha \subseteq V \cap \prod_{\alpha\in\Delta} A_\alpha.$$

It follows that every τ_X-open set containing x meets $A = \prod_{\alpha\in\Delta} A_\alpha$, hence $x \in \tau_X - Cl(A)$. ∎

In the following, we are concerned with necessary conditions and sufficient conditions for the M-continuity of a function on a space with a minimal structure into a product space of a family of spaces with minimal structure.

We consider a family of spaces with minimal structure $\widetilde{\mathcal{F}} = \{(Y_\alpha, \widetilde{m}_\alpha) : \alpha \in \Delta\}$ and the cartesian product $Y := \prod\{Y_\alpha : \alpha \in \Delta\}$. For each $\alpha \in \Delta$ we denote by $P_a : Y \to Y_\alpha$ the canonical projection.

THEOREM 4.4. *Let (X, m_X) be a space with minimal structure and let $f : X \to Y$ be defined by $f(x) = \{f_\alpha(x)\}_{\alpha\in\Delta}$, $x \in X$. If m_Y is a mini-*

mal structure on Y such that $f:(X,m_X) \to (Y,m_Y)$ is M-continuous and if $P_\alpha : (Y,m_Y) \to (Y_\alpha,\widetilde{m}_\alpha)$ is M-continuous for some $\alpha \in \Delta$, then $f_\alpha : (X,m_X) \to (Y_\alpha,\widetilde{m}_\alpha)$ is M-continuous.

PROOF. We notice that $f_\alpha = P_\alpha \circ f$ and apply Corollary 2.2. ∎

COROLLARY 4.1. *Let (X,m_X), (Y,m_Y) and $(X \times Y, m_{X\times Y})$ be spaces with minimal structure and let $F: X \to Y$ be a function. If the graph function $g:(X,m_X) \to (X \times Y, m_{X\times Y})$, defined by $g(x) = (x,F(x))$, $x \in X$, and the canonical projection $P_2 : (X \times Y, m_{X\times Y}) \to (Y,m_Y)$ are M-continuous, then $F:(X,m_X) \to (Y,m_Y)$ is M-continuous.*

PROOF. We apply Theorem 4.4 in the following context: $\Delta = \{1,2\}$, $Y_1 := X$, $Y_2 =: Y$, $Y := Y_1 \times Y_2$, $\widetilde{m}_1 := m_X$, $\widetilde{m}_2 := m_Y$, $m_Y := m_{X\times Y}$, $f_1(x) = x$, $f_2(x) = F(x)$ for every $x \in X$, $f := g$ and $P_\alpha = P_2$. ∎

THEOREM 4.5. *Let (X,m_X) be a space with minimal structure. Let $\widetilde{\mathcal{F}} = \{(Y_\alpha,\widetilde{m}_\alpha) : \alpha \in \Delta\}$ be a family of spaces with minimal structure and let τ_Y be the topology determined by $\widetilde{\mathcal{F}}$ on Y. Let $f: X \to Y$ be defined by $f(x) = \{f_\alpha(x)\}_{a\in\triangle}$, $x \in X$.*

a) If $f:(X,m_X) \to (Y,\tau_Y)$ is M-continuous, then $f_\alpha : (X,m_X) \to (Y_\alpha,\widetilde{m}_\alpha)$ is M-continuous for each $\alpha \in \Delta$;

b) If f_α is M-continuous for each $\alpha \in \Delta$ and m_X is a base for a topology on X, then f is M-continuous.

PROOF. a) This follows by Theorem 4.4, since, by the definition of τ_Y, the projection $P_\alpha : (Y,\tau_Y) \to (Y_\alpha,\widetilde{m}_\alpha)$ is M-continuous for each $\alpha \in \Delta$.

b) Assume that m_X is a base for a topology on X. Let $V \in \mathcal{S}$. There exists $\alpha \in \Delta$ and $U_\alpha \in \widetilde{m}_\alpha$ such that $V = P_\alpha^{-1}(U_\alpha)$. Then $f^{-1}(V) = (P_\alpha \circ f)^{-1}(U_\alpha) = f_\alpha^{-1}(U_\alpha)$. Since $f_\alpha : (Y,m_Y) \to (Y_\alpha,\widetilde{m}_\alpha)$ is M-continuous, it follows that $f^{-1}(V) \in \mathcal{U}(m_X)$. Then $f^{-1}(\mathcal{S}) \subseteq \mathcal{U}(m_X)$, hence $f^{-1}(\mathcal{I}_F(\mathcal{S})) = \mathcal{I}_F(f^{-1}(\mathcal{S})) \subseteq \mathcal{I}_F(\mathcal{U}(m_X))$. By our assumption and by Lemma 2.1, $\mathcal{I}_F(\mathcal{U}(m_X)) \subseteq \mathcal{U}(m_X)$. Then $f^{-1}(\tau_X) = f^{-1}(\mathcal{U}(\mathcal{I}_F(\mathcal{S}))) = \mathcal{U}(f^{-1}(\mathcal{I}_F(\mathcal{S}))) \subseteq \mathcal{U}(\mathcal{U}(m_X)) = \mathcal{U}(m_X)$. By Lemma 2.4, it follows that $f:(X,m_x) \to (Y,\tau_Y)$ is M-continuous. ∎

THEOREM 4.6. *Let (X,m_X) be a space with minimal structure.*

Assume that for every family $\mathcal{F}$ of two spaces with minimal structure, $\mathcal{F} = \{(Y_1,\widetilde{m}_1),(Y_2,\widetilde{m}_2)\}$ which determines on $Y = Y_1 \times Y_2$ a topology τ_Y, *and for every function $f:(X,m_X) \to (Y,\tau_Y)$, $f=(f_1,f_2)$, the M-continuity of f_1 and f_2 implies the M-continuity of f. Then m_X is a base for a topology on X.*

PROOF. Let us take $Y_1 = Y_2 = X$ and $\widetilde{m}_1 = \widetilde{m}_2 = m_X$. For $k=1,2$, let f_k be the identity mapping of X. Obviuously, $f_k : (X,m_X) \to (Y_k,\widetilde{m}_k)$ is M-continuous. Then the function $f:(X,m_X) \to (Y,\tau_Y)$ defined by $f(x) = = (f_1(x),f_2(x)) = (x,x)$, is M-continuous. Take U, $V \in m_X$ such that $U \cap V$ is nonempty. We will prove that $U \cap V \in \mathcal{U}(m_X)$. Then we obtain by induction that $\mathcal{I}_F(m_X) \subseteq \mathcal{U}(m_X)$, hence m_X is a base for a topology on X, by Lemma 3.2. We have $(U \cap V) \times X = P_1^{-1}(U) \cap P_2^{-1}(V) \in \mathcal{I}_F(\mathcal{S}) \subseteq \tau_Y$. Since f is M-continuous, this yields that $U \cap V = f^{-1}((U \cap V) \times X)$ belongs to $\mathcal{U}(m_X)$. ∎

In the following result, we consider two families of spaces with minimal structure with the same set of indices $\mathcal{F} = \{(X_\alpha,m_\alpha) : \alpha \in \Delta\}$ and $\widetilde{\mathcal{F}} = = \{(Y_\alpha,\widetilde{m}_\alpha) : \alpha \in \Delta\}$. Let $X := \prod\{X_\alpha : \alpha \in \Delta\}$ (resp. $Y := \prod\{Y_\alpha : \alpha \in \in \Delta\}$) be endowed with a minimal structure m_X (resp. m_Y). For each $\alpha \in \Delta$, denote by $\pi_\alpha : X \to X_\alpha$, $P_\alpha : Y \to Y_\alpha$ the canonical projections.

THEOREM 4.7. *Assume that, for some $\beta \in \Delta$, $\pi_\beta : (X,m_X) \to (X_\beta,m_\beta)$ is almost M-open and $P_\beta : (Y,m_Y) \to (Y_\beta,\widetilde{m}_\beta)$ is M-continuous. Let $f:X \to Y$ be a product function, defined by $f(\{x_\alpha\}_{\alpha\in\Delta}) = \{f_\alpha(x_\alpha)\}_{\alpha\in\Delta}$, for each $\{x_\alpha\}_{\alpha\in\Delta} \in X$. If $f:(X,m_X) \to (Y,m_Y)$ is M-continuous, then $f_\beta : (X_\beta,m_\beta) \to (Y_\beta,\widetilde{m}_\beta)$ is M-continuous.*

PROOF. We notice that, for each $\alpha \in \Delta$, we have $f_\alpha \circ \pi_\alpha = P_\alpha \circ f$ and $\pi_\alpha(\pi_\alpha^{-1}(A_\alpha)) = A_\alpha$ whenever $A_\alpha \subseteq X_\alpha$. Let $V_\beta \in \widetilde{m}_\beta$. In order to prove that $f_\beta : (X_\beta,m_\beta) \to (Y_\beta,\widetilde{m}_\beta)$ is M-continuous, it is necessary and sufficient to check that $f_\beta^{-1}(V_\beta) \in \mathcal{U}(m_\beta)$, according to Lemma 2.4.

Since $P_\beta : (Y,m_Y) \to (Y_\beta,\widetilde{m}_\beta)$ is M-continuous, $P_\beta^{-1}(V_\beta) \in \mathcal{U}(m_Y)$ by Lemma 2.4. The M-continuity of $f:(X,m_X) \to (Y,m_Y)$ implies, according to Proposition 2.1, $f^{-1}(P_\beta^{-1}(V_\beta)) \in \mathcal{U}(m_X)$. But $f^{-1}(P_\beta^{-1}(V_\beta)) = = \pi_\beta^{-1}(f_\beta^{-1}(V_\beta))$, hence $f_\beta^{-1}(V_\beta) = \pi_\beta(f^{-1}(P_\beta^{-1}(V_\beta))) \in \pi_\beta(\mathcal{U}(m_X)) = \mathcal{U}(\pi_\beta(m_X))$.

Since $\pi_\beta : (X, m_X) \to (X_\beta, m_\beta)$ is almost M-open, we have $\pi_\beta(m_X) \subseteq \subseteq \mathcal{U}(m_\beta)$, by Lemma 2.5. It follows that $f_\beta^{-1}(V_\beta) \in \mathcal{U}(\mathcal{U}(m_\beta)) = \mathcal{U}(m_\beta)$. ∎

5. Some types of minimal structures on a cartesian product

Now we give some applications to the results from the preceding section. We will pay attention to some minimal structures which can be defined on a topological space (X, τ), namely $GO_T(X)$ and $\tau \cap F_\sigma(X)$.

THEOREM 5.1. *Let $\{(X_\alpha, \tau_\alpha) : \alpha \in \Delta\}$ be a family of topological spaces with the product space denoted by (X, τ). Let T be a finite composition of operators Int and Cl. Then each canonical projection $\pi_\alpha : (X, GO_T(X)) \to \to (X_\alpha, GO_T(X_\alpha))$ is M-continuous and M-open.*

PROOF. It is well-known that each projection $\pi_\alpha : (X, \tau) \to (X_\alpha, \tau_\alpha)$ is continuous and open. The claim follows by Theorem 3.1. ∎

COROLLARY 5.1. *Let $A_\alpha \subseteq X_\alpha$, $\alpha \in \Delta$. Then $\pi_\alpha^{-1}(A_\alpha) \in GO_T(X)$ if and only if $A_\alpha \in GO_T(X_\alpha)$.*

PROOF. The necessity follows by the M-openness of

$$\pi_\alpha : (X, GO_T(X)) \to (X_\alpha, GO_T(X_\alpha)),$$

since $A_\alpha = \pi_\alpha(\pi_\alpha^{-1}(A_\alpha))$. The sufficiency follows by the M-continuity of $\pi_\alpha : (X, GO_T(X)) \to (X_\alpha, GO_T(X_\alpha))$, using Lemma 2.4 and taking into account that $GO_T(X)$ has property $(\mathscr{B})$, according to Lemma 3.1. ∎

In the setting of Theorem 5.1, we give a generalization of Corollary 5.1 which unifies many known results.

THEOREM 5.2. *Let $n \geq 1$, $\alpha_1, \alpha_2, \ldots, \alpha_n \in \Delta$ and $A_{\alpha_k} \subseteq X_{\alpha_k}$, $k \in \{1, 2, \ldots, n\}$. $\bigcap_{k=1}^{n} \pi_{\alpha_k}^{-1}(A_{\alpha_k}) \in GO_T(X)$ if and only if $A_{\alpha_k} \in GO_T(X_{\alpha_k})$ for each $k \in \{1, 2, \ldots, n\}$.*

PROOF. Denote $A = \bigcap_{k=1}^{n} \pi_{\alpha_k}^{-1}(A_{\alpha_k}) = A_{\alpha_1} \times A_{\alpha_2} \times \cdots \times A_{\alpha_n} \times \prod_{\beta \neq \alpha_k} X_\beta$.

NECESSITY. For every $\alpha \in \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ we have $\pi_\alpha(A) = A_\alpha$, and the necessity follows by M-openness of π_α.

SUFFICIENCY. It is known that $Cl\Big(\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big) = \bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(Cl(M_{\alpha_k}))$ and $Int\Big(\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big) = \bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(Int(M_{\alpha_k}))$ whenever. $n \geq 1$, $\alpha_1,\alpha_2,\ldots$ $\ldots,\alpha_n \in \Delta$ and $M_{\alpha_k} \subseteq X_{\alpha_k}$, $k \in \{1,2,\ldots,n\}$. The first equality above is a particular case of $Cl\Big(\prod\limits_{\alpha\in\Delta} M_\alpha\Big) = \prod\limits_{\alpha\in\Delta} Cl(M_\alpha)$. The second equality follows if we notice that $\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(Int(M_{\alpha_k}))$ is an open set in X, contained in $\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})$, that is $\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(Int(M_{\alpha_k})) \subseteq Int\Big(\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big)$, and for every $j \in \{1,2,\ldots,n\}$ we have

$$\pi_{\alpha_j}\left(Int\Big(\bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big)\right) \subseteq Int\left(\pi_{\alpha_j}\left(\bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\right)\right) = Int(M_{\alpha_j}),$$

whence

$$Int\Big(\bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big) \subseteq \bigcap_{j=1}^{n}\pi_{\alpha_j}^{-1}(Int(M_{\alpha_j})).$$

We conclude that $T\Big(\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(M_{\alpha_k})\Big) = \bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(T(M_{\alpha_k}))$ whenever. $n \geq 1$, $\alpha_1,\alpha_2,\ldots,\alpha_n \in \Delta$ and $M_{\alpha_k} \subseteq X_{\alpha_k}$, $k \in \{1,2,\ldots,n\}$. If $A_{\alpha_k} \in GO_T(X_{\alpha_k})$ for each $k \in \{1,2,\ldots,n\}$, then

$$\bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(A_{\alpha_k}) \subseteq \bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(T(A_{\alpha_k})) = T\Big(\bigcap_{k=1}^{n}\pi_{\alpha_k}^{-1}(A_{\alpha_k})\Big),$$

whence $\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(A_{\alpha_k}) \in GO_T(X)$. This proves the sufficiency part. ∎

COROLLARY 5.2. *Let* $n \geq 1$, α_1, α_2, …, $\alpha_n \in \Delta$ *and* $A_{\alpha_k} \subseteq X_{\alpha_k}$, $k \in \{1,2,\ldots,n\}$. *Then* $\bigcap\limits_{k=1}^{n}\pi_{\alpha_k}^{-1}(A_{\alpha_k})$ *is semi-open (resp. preopen,* α*-open,* β*-open) if and only if the sets* A_{α_k}, $k \in \{1,2,\ldots,n\}$, *are semi-open (resp. preopen,* α*-open,* β*-open).*

REMARK 5.1. Corollary 5.2 shows that Theorem 5.2 generalizes Lemma 5.1 from (El-Deeb et al. [10]), Lemma 3.1 from (Chae et al. [8]) and a result from (Noiri [17]).

THEOREM 5.3. *Let (X,τ) be a topological space and let $\{(Y_\alpha,\widetilde{\tau}_\alpha):\alpha \in \in \Delta\}$ be a family of topological spaces with its product space $(Y,\widetilde{\tau})$. Let T and $\widetilde{T}$ be two finite compositions of operators Int and Cl. Let $f\colon X \to \to Y$ be defined by $f(x) = \{f_\alpha(x)\}_{a\in\triangle}$, $x \in X$. If $f\colon(X, GO_T(X)) \to (Y, GO_{\widetilde{T}}(Y))$ is M-continuous, then $f_\alpha : (X, GO_T(X)) \to (Y_\alpha, GO_{\widetilde{T}}(Y_\alpha))$ is M-continuous for each $\alpha \in \Delta$.*

PROOF. Let $\alpha \in \Delta$. By Theorem 5.1, the canonical projection P_α : $(Y, GO_T(Y)) \to (Y_\alpha, GO_{\widetilde{T}}(Y_\alpha))$ is M-continuous. Then we can apply Theorem 4.4 with $m_X = GO(T)$, $m_Y = GO_{\widetilde{T}}(Y)$ and $\widetilde{m}_\alpha = GO_{\widetilde{T}}(Y_\alpha)$. ∎

COROLLARY 5.3. *Let X and Y be two topological spaces and let $F\colon X \to Y$. If the graph function $g\colon(X, GO_T(X)) \to (X\times Y, GO_{\widetilde{T}}(X\times Y))$ defined by $g(x) = (x, F(x))$, $x \in X$, is M-continuous, then $F\colon(X, GO_T(X)) \to \to (Y, GO_{\widetilde{T}}(Y))$ is M-continuous.*

THEOREM 5.4. *Let $\{(X_\alpha,\tau_\alpha) : \alpha \in \Delta\}$ and $\{(Y_\alpha,\widetilde{\tau}_\alpha) : \alpha \in \Delta\}$ be two families of topological spaces with the same set of indices. Let (X,τ), respectively $(Y,\widetilde{\tau})$, be the corresponding product spaces. Let T and $\widetilde{T}$ be two finite compositions of operators Int and Cl. Let $f\colon X \to Y$ be a product function defined by $f(\{x_\alpha\}_{\alpha\in\Delta}) = \{f_\alpha(x_\alpha)\}_{a\in\triangle}$, for each $\{x_\alpha\}_{\alpha\in\Delta} \in X$.*

If $f\colon(X, GO_T(X)) \to (Y, GO_{\widetilde{T}}(Y))$ is M-continuous, then

$$f_\alpha\colon\bigl(X_\alpha, GO_T(X_\alpha)\bigr) \to (Y_\alpha, GO_{\widetilde{T}}(Y_\alpha))$$

is M-continuous for each $\alpha \in \Delta$.

PROOF. Let $\alpha \in \Delta$. Denote by $\pi_\alpha : X \to X_\alpha$ and $P_\alpha : Y \to Y_\alpha$, $\alpha \in \Delta$, the canonical projections. By Theorem 5.1, $\pi_\alpha : (X, GO_T(X)) \to \to (X_\alpha, GO_T(X_\alpha))$ is M-open and $P_\alpha : (Y, GO_{\widetilde{T}}(Y)) \to (Y_\alpha, GO_{\widetilde{T}}(Y_\alpha))$ is M-continuous. Then we can apply Theorem 4.7 with $\beta = \alpha$, $m_X = GO_T(X)$, $m_Y = GO_{\widetilde{T}}(Y)$, $m_\alpha = GO_T(X_\alpha)$ and $\widetilde{m}_\alpha = GO_{\widetilde{T}}(Y_\alpha)$. ∎

COROLLARY 5.4. *$f\colon(X, GO_T(X)) \to (Y,\widetilde{\tau})$ is M-continuous if and only if $f_\alpha : (X_\alpha, GO_T(X_\alpha)) \to (Y_\alpha,\widetilde{\tau}_\alpha)$ is M-continuous for each $\alpha \in \Delta$.*

PROOF. Let us take in the statement of Theorem 5.4: $\widetilde{T} = Int$, i.e.

$$GO_{\widetilde{T}}(Y) = \widetilde{\tau} \quad \text{and} \quad GO_{\widetilde{T}}(Y_\alpha) = \widetilde{\tau}_\alpha.$$

NECESSITY. It follows from Theorem 5.4.

SUFFICIENCY. Let V be a basic open set in Y, that is $V = \prod_{j=1}^{n} V_{\alpha_j} \times \times \prod_{\alpha \neq \alpha_j} Y_\alpha$, where $n \geq 1$, $\{\alpha_j : j = 1, \ldots, n\} \subseteq \Delta$, and V_{α_j} is an open set in X_{α_j} for each $j \in \{1, \ldots, n\}$. We have $f^{-1}(V) = \prod_{j=1}^{n} f_{\alpha_j}^{-1}(V_{\alpha_j}) \times \prod_{\alpha \neq \alpha_j} X_\alpha$. By our assumption, $f_{\alpha_j}^{-1}(V_{\alpha_j}) \in GO_T(X_{\alpha_j})$ for each $j \in \{1, \ldots, n\}$. It follows that $f^{-1}(V) \in GO_T(X)$, by Theorem 5.2. Let W be an arbitrary open set in Y. Then there exist a family $\{V_i : i \in I\}$ of basic open sets in Y such that $W = \bigcup_{i \in I} V_i$. Then $f^{-1}(W) = \bigcup_{i \in I} f^{-1}(V_i) \in \mathcal{U}(GO_T(X))(= GO_T(X))$. It follows that $f: (X, GO_T(X)) \to (Y, \widetilde{\tau})$ is M-continuous. ∎

REMARK 5.2. Let X and Y be two topological spaces.

1) Let T and $\widetilde{T}$ be two finite compositions of operators Int and Cl. An M-continuous function $f: (X, GO_T(X)) \to (Y, GO_{\widetilde{T}}(Y))$ is called:

i) strongly α-continuous if $T = Int\,Cl\,Int$ and $\widetilde{T} = Cl\,Int$ (Beceren, [3]);

ii) semi$-\alpha$-irresolute if $T = Cl\,Int$ and $\widetilde{T} = Int\,Cl\,Int$ (Beceren, [4]);

iii) almost α-irresolute if $T = Cl\,Int\,Cl$ and $\widetilde{T} = Int\,Cl\,Int$ (Beceren, [5]);

iv) α-precontinuous if $T = Int\,Cl$ and $T = Int\,Cl\,Int$ (Beceren, [6]);

v) α-preirresolute (resp. β-preirresolute) if $T = Int\,Cl\,Int$ (resp. $T = = Cl\,Int\,Cl$) and $\widetilde{T} = Int\,Cl$ (Beceren and Noiri, [7]).

This observation shows that:

a) Proposition 4.3 generalizes Theorems 3.3 from [3], [4], [5], [6] and Theorem 4.3 from [7].

b) Corollary 4.3 generalizes Theorems 3.2 from [3], [4], [5], [6] and Theorem 4.2 from [7].

c) Proposition 4.4 generalizes Theorems 3.4 from[3], [4], [5], [6] and Theorem 4.4 from [7].

2) Let (X,τ) and $(Y,\widetilde{\tau})$ be topological spaces. A function $f: X \to Y$ is said to be H-almost continuous at $x \in X$ if for each open set $V \subseteq Y$ containing $f(x)$, the closure of $f^{-1}(V)$ is a neighborhood of x. A function $f: X \to Y$ is said to be H-almost continuous if it is H-almost continuous at every point $x \in X$. $f: X \to Y$ is H-almost continuous if and only if M-continuous function $f: (X, PO(X)) \to (Y,\widetilde{\tau})$ is M-continuous, by Theorem 1 from (Popa [20]).

Theorem 5.3 generalizes Theorem 6 from (Anderson and Jensen [1]). Corollary 5.4 generalizes Theorem 5 from (Popa [20]).

Now we show that some results from the preceding section generalize known results concerning D-continuous functions and D-supercontinuous functions (see Remark 2.2 and Remark 3.4).

COROLLARY 5.5. (Kohli [11], Theorem 2.10]) *Let $\{(Y_\alpha,\widetilde{\tau}_\alpha) : \alpha \in \Delta\}$ be a family of topological spaces with the product space denoted by $(Y,\widetilde{\tau})$, and let (X,τ) be a topological space. If a function $f: X \to Y$, defined by $f(x) = \{f_\alpha(x)\}_{\alpha\in\Delta}$, $x \in X$, is D-continuous, then f_α is D-continuous for each $\alpha \in \Delta$.*

PROOF. The function $f: (X,\tau) \to (Y,\widetilde{\tau} \cap F_\sigma(Y))$ is M-continuous by our assumption. Let $\alpha \in \Delta$. The canonical projection $P_\alpha : (Y,\widetilde{\tau}) \to (Y_\alpha,\widetilde{\tau}_\alpha)$ is continuous, hence $P_\alpha : (Y,\widetilde{\tau} \cap F_\sigma(Y)) \to (Y_\alpha,\widetilde{\tau}_\alpha \cap F_\sigma(Y_\alpha))$ is M-continuous (see Example 2.1). It follows by Theorem 4.4 that $f_\alpha : (X,\tau) \to (Y_\alpha,\widetilde{\tau}_\alpha \cap$ $\cap F_\sigma(Y_\alpha))$ is M-continuous. ∎

COROLLARY 5.6. (Kohli [11], Theorem 2.4) *Let (X,τ) and $(Y,\widetilde{\tau})$ be topological spaces and let $F: X \to Y$. If the graph function $g: X \to X \times Y$ is D-continuous, then F is D-continuous.*

COROLLARY 5.7. (Kohli [11], Theorem 2.9) *Let $\{(X_\alpha,\tau_\alpha) : \alpha \in \Delta\}$ and $\{(Y_\alpha,\widetilde{\tau}_\alpha) : \alpha \in \Delta\}$ be two families of topological spaces with the same set of indices. Let (X,τ), respectively $(Y,\widetilde{\tau})$, be the corresponding product spaces. Let $f: X \to Y$ be a product function defined by $f(\{x_\alpha\}_{\alpha\in\Delta}) = \{f_\alpha(x_\alpha)\}_{a\in\triangle}$, for each $\{x_\alpha\}_{\alpha\in\Delta} \in X$. If f is D-continuous, then each f_α is D-continuous.*

PROOF. The function $f: (X,\tau) \to (Y,\widetilde{\tau} \cap F_\sigma(Y))$ is M-continuous by our assumption. Let $\alpha \in \Delta$. Denote by $\pi_\alpha : X \to X_\alpha$ and $P_\alpha : Y \to Y_\alpha$, $\alpha \in \Delta$, the canonical projections. Since π_α is open and $P_\alpha : (Y,\widetilde{\tau} \cap F_\sigma(Y)) \to$ $\to (Y_\alpha,\widetilde{\tau}_\alpha \cap F_\sigma(Y_\alpha))$ is M-continuous, we can apply Theorem 4.7. It follows that $f_\alpha : (X_\alpha,\tau_\alpha) \to (Y_\alpha,\widetilde{\tau}_\alpha \cap F_\sigma(Y_\alpha))$ is M-continuous. ∎

COROLLARY 5.8. (Kohli and Singh [12], Theorem 3.8) *Let* $\{(Y_\alpha, \widetilde{\tau}_\alpha) : \alpha \in \Delta\}$ *be a family of topological spaces with the product space denoted by* $(Y, \widetilde{\tau})$, *and let* (X, τ) *be a topological space. If a function* $f: X \to Y$, *defined by* $f(x) = \{f_\alpha(x)\}_{\alpha \in \Delta}$, $x \in X$, *is* D*-supercontinuous, then* f_α *is* D*-supercontinuous for each* $\alpha \in \Delta$.

PROOF. Since $f: (X, \tau \cap F_\sigma(X)) \to (Y, \widetilde{\tau})$ is M-continuous and $P_\alpha : (Y, \widetilde{\tau}) \to (Y_\alpha, \widetilde{\tau}_\alpha)$ is continuous, it follows that $f_\alpha : (X, \tau \cap F_\sigma(X)) \to$ $\to (Y_\alpha, \widetilde{\tau}_\alpha)$ is M-continuous. ∎

A topological space (X, τ) is said to be D-regular if it has a base consisting of open F_σ-sets (Kohli [11], Kohli and Singh [12]); this means that $\tau \subseteq \mathcal{U}(\tau \cap F_\sigma(X))$. It is easy to see that every continuous function on a D-regular space is D-supercontinuous. ∎

COROLLARY 5.9. (Kohli and Singh [12], Theorem 3.9) *Let* (X, τ) *and* $(Y, \widetilde{\tau})$ *be topological spaces and let* $F: X \to Y$. *Then the graph function* $g: X \to X \times Y$ *is* D*-supercontinuous if and only if* F *is* D*-supercontinuous and* X *is* D*-regular.*

PROOF. NECESSITY. Since $g: X \to X \times Y$ is D-supercontinuous and $F = P_2 \circ g$, where $P_2 : X \to X \times Y$ is the canonical projection, F is D-supercontinuous, by Corollary 5.8. Let $U \in \tau$. Then $U \times Y$ is open in $X \times Y$, hence $U = g^{-1}(U \times Y) \in \mathcal{U}(\tau \cap F_\sigma(X))$, by Lemma 2.4. We have proven that $\tau \subseteq \mathcal{U}(\tau \cap F_\sigma(X))$, hence X is D-regular.

SUFFICIENCY. Since X is D-regular by our assumption, it suffices to prove that g is continuous, i.e. $g^{-1}(W) \in \tau$ whenever $W \subseteq X \times Y$ is an open set. There exist some families of open sets, $\{U_i : i \in\} \subseteq \tau$ and $\{V_i : i \in I\} \subseteq \widetilde{\tau}$, such that $W = \bigcup\{U_i \times V_i : i \in I\}$. Since F is continuous, $F^{-1}(V_i) \in \tau$ for every $i \in I$. Then $g^{-1}(W) = \bigcup\{g^{-1}(U_i \times V_i) : i \in I\} =$ $= \bigcup\{U_i \cap F^{-1}(V_i) : i \in I\} \in \tau$. ∎

References

[1] D. R. ANDERSON and J. A. JENSEN, Semi-continuity on topological spaces, *Atti Acad. Naz. Lincei, Rend. Cl. Sci. Math. Fiz. Natur.* **181(42)** (1973), 782–785.

[2] M. E. ABD EL-MONSEF, S. N. EL-DEEB and R. A. MAHMOUD, β-open sets and β-continuous mappings, *Bull. Fac. Sci. Assiut Univ.* **12** (1983), 77–90.

[3] Y. BECEREN, On strongly α-continuous functions, *Far East J. Math. Sci. (FJMS)* **Special Volume** (2000), Part I (Geometry and Topology), 51–58.

[4] Y. BECEREN, On semi α-irresolute functions, *J. Indian Acad. Math.* **22(2)** (2000), 353–362.

[5] Y. BECEREN, Almost α-irresolute functions, *Bull. Bal. Math. Soc.* **92(3)** (2000), 213–218.

[6] Y. BECEREN and T. NOIRI, On α-precontinuous functions, *Far East J. Math. Sci. (FJMS)* **Special Volume** (2000), Part III (Geometry and Topology), 295–303.

[7] Y. BECEREN and T. NOIRI, α-preirresolute and β-preirresolute functions, *Demonstratio Mathematica* **XXXVI(1)** (2003), 208–213.

[8] G. I. CHAE, T. NOIRI and D. W. LEE, On na-continuous functions, *Kyungpook Math. J.* **26(1)** (1986), 73–79.

[9] Á. CSÁSZÁR, Generalized topology, generalized continuity, *Acta Math. Hungar.* **96** (2002), 351–357.

[10] N. EL-DEEB, I. A. HASANEIN, A. S. MASHHUR and T. NOIRI, On P-regular spaces, *Bull. Math. Soc. Sci. Math. R. S. R.* **27(75)4** (1983), 311–315.

[11] J. K. KOHLI, D-continuous functions, D-regular spaces and D-Hausdorff spaces; *Bull. Cal. Math. Soc.* **84** (1992), 39–46.

[12] J. K. KOHLI and D. SINGH, D-supercontinuous functions, *Indian J. Pure Appl. Math.* **32(2)** (2001), 227–235.

[13] N. LEVINE, Semi-open sets and semicontinuity in topological spaces, *Amer. Math. Monthly* **70** (1963), 36–41.

[14] A. S. MASHHOUR, M. E. ABD EL-MONSEF and S. N. EL-DEEB, On precontinuous and weak precontinuous mappings, *Proc. Math. Phys. Soc. Egypt* **53** (1982), 47–53.

[15] A. S. MASHHOUR, I. A. HASANEIN and S. N. EL-DEEB, α-continuous and α-open mappings, *Acta Math. Hung.* **41(3–4)** (1983), 213–218.

[16] O. NJÅSTAD, On some classes of nearly open sets, *Pacific J. Math.* **15** (1965), 961–970.

[17] T. NOIRI, Remarks on semi-open mappings, *Bull. Cal. Math. Soc.* **65** (1973), 197–201.

[18] T. NOIRI and V. POPA, A unified theory of contra-continuity for functions, *Annale Univ. Sci. Budapest* **44** (2002), 115–137.

[19] T. NOIRI and V. POPA, A unified theory of θ-continuity for functions, *Rendiconti del Circolo Matematico di Palermo*, Serie II, Tomo LII (2003), pp. 163–188.

[20] T. NOIRI and V. POPA, Faintly m-continuous functions; *Chaos, Solitons and Fractals* **19** (2004), 1147–1159.

[21] V. POPA, Properties of H-almost continuous functions, *Bull. Math. de Soc. Math. de la R. S. de Roumanie*, **Tome 31 (79), nr. 2** (1987).

[22] V. POPA and T. NOIRI, On M-continuous functions, *Anal. Univ. "Dunarea de Jos" Galati, Ser. Mat. Fiz. Mec. Teor.* **Fasc. II, 18(23)** (2000), 31–41.

[23] V. POPA and T. NOIRI, On the definitions of some generalized forms of continuity under minimal conditions, *Mem. Fac. Sci., Kochi Univ. (Math.)* **22** (2001), 9–18.

[24] V. POPA and T. NOIRI, Slightly m-continuous functions, *Radovi Matematički* **10** (2001), 219–232.

[25] V. POPA and T. NOIRI, On weakly (τ, m)-continuous functions, *Rendiconti del Circolo Matematico di Palermo*, Serie II, Tomo LI (2002), pp. 295–316.

[26] V. POPA and T. NOIRI, A unified theory of weak continuity for functions, *Rendiconti del Circolo Matematico di Palermo*, Serie II, Tomo LI (2002), pp. 439–464.

[27] N. V. VELIČKO, H-closed topological spaces, *Am. Math. Soc. Transl.* **78(2)** (1968), 103–118.

Marcelina Mocanu

Department of Mathematics
University of Bacău
Spiru Haret 8, 600114
Bacău, Romania
marcim@ub.ro

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 91–96

EXTREMALLY DISCONNECTED GENERALIZED TOPOLOGIES

By

ÁKOS CSÁSZÁR*

(Received September 22, 2004)

1. Introduction

According to [6], a *generalized topology* on a set X is a subset μ of the power set $\exp X$ of X such that $\emptyset \in \mu$ and an arbitrary union of elements of μ belongs to μ. For the sake of brevity, generalized topology will be abbreviated by GT. The elements of the GT μ will be called μ*-open*, their complements μ*-closed*.

If $X \in \mu$ then the GT μ will be called a *strong* GT. Of course, a topology is a strong GT such that the intersection of two μ-open sets is always μ-open.

According to [9], if μ is a GT on X and $A \subset X$, then

$$i_\mu A = \bigcup\{M \in \mu : M \subset A\} \tag{1}$$

is a mapping $i_\mu : \exp X \to \exp X$ such that it is monotone, idempotent and restricting, where $\gamma\colon \exp X \to \exp X$ is said to be *monotone* iff $A \subset B \subset X$ implies $\gamma A \subset \gamma B$, *idempotent* iff $\gamma\gamma A = \gamma A$ for $A \subset X$, *restricting* iff $\gamma A \subset A$ for $A \subset X$.

Similarly, if

$$c_\mu A = \bigcap\{N : A \subset N, X - N \in \mu\}, \tag{2}$$

then c_μ is again monotone and idempotent, but enlarging, where $\gamma\colon \exp X \to$ $\to \exp X$ is said to be *enlarging* iff $A \subset \gamma A$ for $A \subset X$. Further each of i_μ and c_μ determines the other one as $A \subset X$ implies

$$c_\mu A = X - i_\mu(X - A). \tag{3}$$

* Research supported by Hungarian Foundation for Scientific Research, grant No. T 032042.

If μ is a topology then clearly $i_\mu A = \mathrm{int}(A)$ and $c_\mu A = \mathrm{cl}(A)$ for $A \subset X$.

Clearly each set $i_\mu A$ is μ-open and $c_\mu A$ is μ-closed.

It is well-known that, in the literature, a topological space or a topology is called *extremally disconnected* (briefly EDC) iff the closure of an open set is always open. Similarly, we shall say that a GT μ is EDC iff $c_\mu A \in \mu$ whenever $A \in \mu$.

In the following, we shall present some interesting examples of EDC GT's.

2. Extremal disconnectedness of $\alpha, \sigma, \pi, \beta$

Let us denote by Γ the collection of all monotone mappings $\gamma : \exp X \to \to \exp X$. By [4], if $\gamma \in \Gamma$, the sets A satisfying $A \subset \gamma A$ constitute a GT, denoted by λ_γ. The elements of λ_γ are said to be (λ_γ-open or) γ-*open*, their complements (λ_γ-closed or) γ-*closed*.

By [4], 1.8, a set A is γ-closed iff $A \supset \gamma^* A$, where

$$\gamma^* A = X - \gamma(X - A) \quad (A \subset X) \tag{4}$$

is the *conjugate* of γ; clearly $\gamma \in \Gamma$ implies $\gamma^* \in \Gamma$ and $(\gamma^*)^* = \gamma$.

If $\gamma, \gamma' \in \Gamma$, we write for the sake of brevity $\gamma\gamma'$ instead of $\gamma \circ \gamma'$. Clearly $(\gamma\gamma')^* = \gamma^*(\gamma')^*$.

We recall that, in [9], some operations were generalized for arbitrary GT's. More precisely, if μ is a GT on X, we define GT's $\alpha(\mu) = \lambda_{i_\mu c_\mu i_\mu}$, $\sigma(\mu) = \lambda_{c_\mu i_\mu}$, $\pi(\mu) = \lambda_{i_\mu c_\mu}$, $\beta(\mu) = \lambda_{c_\mu i_\mu c_\mu}$.

In the literature, in the case when μ is a topology, the elements of $\alpha(\mu)$ are called α-*open* [12], those of $\sigma(\mu)$ *semi-open* [10], those of $\pi(\mu)$ *preopen* [11], those of $\beta(\mu)$ β-*open* [1].

In the following, if μ is a fixed GT, we shall simply write $\alpha, \sigma, \pi, \beta$ instead of $\alpha(\mu), \sigma(\mu), \pi(\mu), \beta(\mu)$. Moreover, we write ι for i_μ and κ for c_μ.

We recall the inclusions

$$\mu \subset \alpha \subset \sigma \subset \beta, \qquad \alpha \subset \pi \subset \beta. \tag{5}$$

(See [9], 2.1.)

THEOREM 2.1 *For an arbitrary GT μ on X, if $\gamma \in \Gamma$ is monotone and enlarging then $\nu = \lambda_{\kappa\iota\gamma}$ is EDC.*

PROOF. For $A \subset X$, $A \in \mu$ implies $\iota A = A \subset \gamma A$, hence $\iota A \in \mu$ and $\iota A \subset \gamma A$ so that $\iota A \subset \iota\gamma A$. Therefore, for $A \in \mu$, we have $A = \iota A \subset \iota\gamma A \subset$ $\subset \kappa\iota\gamma A$. We obtain $\mu \subset \nu = \lambda_{\kappa\iota\gamma}$, consequently the μ-closed set $\kappa\iota\gamma A$ is ν-closed. If A is ν-open then $A \subset \kappa\iota\gamma A$ and the latter is ν-closed, hence $c_\nu A \subset \kappa\iota\gamma A$ and $A \subset c_\nu A \subset \kappa\iota\gamma A \subset \kappa\iota\gamma c_\nu A$ so that $c_\nu A \in \nu$ as stated. ∎

THEOREM 2.2 *For an arbitrary GT μ on X, the GT's σ and β are EDC.*

PROOF. We apply 2.1 for $\gamma = \mathrm{id}$ and $\gamma = \kappa$, respectively. ∎

It is interesting to observe that α and π need not be EDC even if μ is a topology. Consider $X = \mathbb{R}$ and let μ be the usual (Euclidean) topology on $\mathbb{R}$. Then $A = (0, 1)$ is open, hence α-open and π-open by (5). The set $[0, 1]$ is closed, hence α-closed and π-closed, and, more precisely, $[0, 1] = c_\alpha A =$ $= c_\pi A$; in fact, the sets $[0, 1)$ and $(0, 1]$ are not π-closed as e.g. $c_\mu i_\mu [0, 1) =$ $= [0, 1]$ is not contained in $[0, 1)$. Consequently these sets are not α-closed either. Now $[0, 1]$ is not π-open (hence not α-open) since $i_\mu c_\mu [0, 1] = (0, 1)$.

3. Unions and intersections of monotone mappings

In order to obtain a further interesting example of EDC GT's, we need some preparatory work.

Let again Γ be the collection of monotone mappings on $\exp X$ and I a non-empty index set. Suppose $\gamma_i \in \Gamma$ for $i \in I$. Let us define $\varphi : \exp X \to$ $\to \exp X$ and $\psi : \exp X \to \exp X$ by

(3.1) $$\varphi A = \bigcup_{i \in I} \gamma_i A$$

and

(3.2) $$\psi A = \bigcap_{i \in I} \gamma_i A$$

for $A \subset X$.

PROPOSITION 3.1 *Both φ and ψ belong to Γ.*

PROOF. If $A \subset B \subset X$, clearly $\gamma_i A \subset \gamma_i B$ for each $i \in I$ so that

$$\varphi A = \bigcup \gamma_i A \subset \bigcup \gamma_i B = \varphi B$$

and

$$\psi A = \bigcap \gamma_i A \subset \bigcap \gamma_i B = \psi B. \quad ∎$$

We write simply

$$\varphi = \bigcup_{i \in I} \gamma_i, \quad \psi = \bigcap_{i \in I} \gamma_i. \tag{3.3}$$

THEOREM 3.2 *With the notation (3.3),*

$$\varphi^* = \bigcap_{i \in I} (\gamma_i)^*, \quad \psi^* = \bigcup_{i \in I} (\gamma_i)^*.$$

PROOF. $\varphi^* A = X - \varphi(X - A) = X - \bigcup \gamma_i (X - A) = \bigcap (X - \gamma_i (X - A)) = = \bigcap (\gamma_i)^* A$. The second equality is proved similarly (we interchange the roles of $\cup$ and $\cap$). ∎

4. i-friendly mappings

Let us consider again a set X and the corresponding collection Γ of monotone mappings. The paper [5] shows that many elements $\gamma \in \Gamma$ have the property

$$i_\mu A = A \cap \gamma A \qquad (A \subset X, \ \mu = \lambda_\gamma). \tag{4.1}$$

By [5] 1.1, (4.1) holds iff $A \cap \gamma A$ is γ-open. Let us say that $\gamma \in \Gamma$ is *i-friendly* iff (4.1) is satisfied, i.e. iff $A \cap \gamma A$ is γ-open for every $A \subset X$.

Observe that, if γ is i-friendly, then by [5], 3.1, $c_\mu A = A \cup \gamma^* A$ for $A \subset X$ and $\mu = \lambda_\gamma$.

According to [5], if μ is a topology, each of the mappings $c_\mu i_\mu$, $i_\mu c_\mu$, $i_\mu c_\mu i_\mu$, $c_\mu i_\mu c_\mu$ is i-friendly. This is not always true if, more generally, μ is a GT. E.g. Example [9], 3.2 shows that there exists a GT μ for which $\gamma = i_\mu c_\mu$ does not fulfil (4.1). However, the following general statement is valid:

THEOREM 4.1 *For an arbitrary GT μ, the mapping $c_\mu i_\mu$ is i-friendly.*

PROOF. By using again the notation $i_\mu = \iota$, $c_\mu = \kappa$, we have, for $A \subset X$, $\iota(A \cap \kappa\iota A) \subset \iota A \subset A \cap \kappa\iota A$ so that $\iota A \in \mu$ implies $\iota A \subset \iota(A \cap \kappa\iota A)$ and $\iota(A \cap \kappa\iota A) = \iota A$. Hence $\kappa\iota(A \cap \kappa\iota A) = \kappa\iota A \supset A \cap \kappa\iota A$. ∎

As $(\kappa\iota)^* = \iota\kappa$, in general, γ can be i-friendly without γ^* being i-friendly ([9], 3.2).

The papers [5] and [8] contain some sufficient conditions for a $\gamma \in \Gamma$ being i-friendly. Let us consider a further such condition:

THEOREM 4.2 *If* γ_i *is* i*-friendly for* $i \in I$ *then* $\varphi = \bigcup_{i \in I} \gamma_i$ *is* i*-friendly.*

PROOF. For $A \subset X$, we have $A \cap \varphi A = A \cap \bigcup \gamma_i A = \bigcup (A \cap \gamma_i A) \subset \subset \bigcup \gamma_i (A \cap \gamma_i A) \subset \bigcup \gamma_i (A \cap \varphi A) = \varphi(A \cap \varphi A)$. ∎

5. Another EDC GT

Now we can prove a statement furnishing further examples of EDC GT's.

THEOREM 5.1 *Let* $\gamma \in \Gamma$ *be such that both* γ *and* γ^* *are* i*-friendly. Then the GT* λ_φ *is EDC where* $\varphi = \gamma \cup \gamma^*$.

PROOF. By 3.1 $\varphi \in \Gamma$. As by 3.2 $\varphi^* = (\gamma \cup \gamma^*)^* = \gamma^* \cap ((\gamma)^*)^* = \gamma^* \cap \gamma$, we have $c_\varphi A = A \cup (\gamma A \cap \gamma^* A)$ for $A \subset X$ because φ is i-friendly by 4.2. Now if $A \subset \varphi A$ then clearly $c_\varphi A \subset \gamma A \cup \gamma^* A \subset \gamma c_\varphi A \cup \gamma^* c_\varphi A = \varphi c_\varphi A$. ∎

COROLLARY 5.2 *If* μ *is a topology and* $\zeta A = c_\mu i_\mu A \cup i_\mu c_\mu A$ *for* $A \subset X$ *then* λ_ζ *is EDC.*

PROOF. Let $\gamma A = c_\mu i_\mu A$ and apply 5.1. ∎

The statement 5.2 is contained in [13], Theorem 3.1. The ζ-open sets are called b-open in [3].

6. Connected EDC GT's

Let us recall that, according to [7], a GT μ (or the space (X, μ)) is said to be connected iff $X = M \cup N$, $M \cap N = \emptyset$, $M, N \in \mu$ imply $M = \emptyset$ or $N = \emptyset$.

If the GT μ is EDC then the connectedness of μ can be very easily characterized:

THEOREM 6.1. *An EDC GT* μ *is connected iff* $M, N \in \mu$, $M \neq \emptyset$, $N \neq \emptyset$ *implies* $M \cap N \neq \emptyset$.

PROOF. The condition is clearly sufficient, even if the GT μ is not EDC. Conversely, if we assume the existence of $M, N \in \mu$ such that $M \neq \emptyset \neq N$ and $M \cap N = \emptyset$ then $\emptyset \neq M \subset c_\mu M \subset X - N \neq X$ and $c_\mu M \in \mu$ since μ is EDC, so that $X = c_\mu M \cup (X - c_\mu M)$, the members are nonempty and disjoint and both belong to μ: μ is not connected. Therefore the condition is sufficient. ∎

A particular case of 6.1 is contained in [2].

References

[1] M. E. ABD EL-MONSEF, S. N. EL-DEEB and R. A. MAHMOUD, β-open sets and β-continuous mappings, *Bull. Fac. Sci. Assiut Univ.* **12** (1966), 77–90.

[2] T. AHO and T. NIEMINEN, Spaces in which preopen subsets are semiopen, *Ricerche Mat.,* **43** (1994), 45–59.

[3] D. ANDRIJEVIĆ, On b-open sets, *Mat. Vesnik* **48** (1996), 59–64.

[4] Á. CSÁSZÁR, Generalized open sets, *Acta Math. Hungar.* **75** (1997), 65–87.

[5] Á. CSÁSZÁR, On the γ-interior and γ-closure of a set, *Acta Math. Hungar.* **80** (1998), 89–93.

[6] Á. CSÁSZÁR, Generalized topology, generalized continuity; *Acta Math. Hungar.* **96** (2002), 351–357.

[7] Á. CSÁSZÁR, γ-connected sets, *Acta Math. Hungar.,* **101** (2003), 273–279.

[8] Á. CSÁSZÁR, Remarks to the formula of γ-interior, *J. Math. Analysis Applic.,* **297** (2004), 490–494.

[9] Á. CSÁSZÁR, Generalized open sets in generalized topologies, *Acta Math. Hungar.,* **106** (2005), 53–66.

[10] N. LEVINE, Semi-open sets and semi-continuity in topological spaces, *Amer. Math. Monthly* **70** (1963), 36–41.

[11] A. S. MASHHOUR, M. E. ABD EL-MONSEF and S. N. EL-DEEB, On precontinuous and weak precontinuous mappings, *Proc. Math. Phys. Soc. Egypt* **53** (1982), 47–53.

[12] O. NJÅSTAD, On some classes of nearly open sets, *Pacific J. Math.* **15** (1965), 961–970.

[13] J. H. PARK, Strongly θ-b-continuous functions, *Acta Math. Hungar.*

Ákos Császár

Department of Analysis,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
csaszar@ludens.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 97–105

NOTES ON SMALL SEMIOVALS

By

GY. KISS* and J. RUFF†

(Received 19 October, 2004)

1. Introduction

Let Π be a projective plane of order q. A *semioval* in Π is a non-empty pointset S with the property that for every point in S there exists a unique line t_P such that $S \cap t_P = \{P\}$. This line is called the tangent to S at P. The classical examples of semiovals arise from polarities (ovals and unitals), and from the theory of blocking sets (the vertexless triangle). The study of semiovals is motivated by their applications to cryptography [1].

It is known that $q + 1 \leq |S| \leq q\sqrt{q} + 1$ and both bounds are sharp [14], [9]. A semioval is said to be *regular with character a* if all nontangent lines intersect S in either 0 or a points. Regular semiovals were studied by Blokhuis and Szőnyi [4], and Gács [7], who proved that in $PG(2,q)$ each regular semioval is either an oval or a unital.

Semiovals with large collinear subsets were investigated by Dover [6]. He proved the following properties of the semioval S:

- $|S \cap \ell| \leq q - 1$ for any line ℓ of Π.
- If S has a $(q-1)$-secant, then $2q - 2 \leq |S| \leq 3q - 3$.
- If S has more than one $(q-1)$-secant, then S can be obtained from a vertexless triangle by removing some subset of points from one side.

* The research was supported by the Hungarian National Foundation for Scientific Research, Grant Nos. T 043556 and T 043758, and by the Slovenian-Hungarian Intergovernmental Scientific and Technological Cooperation Project, Grant No. SLO-1/03.

† The research was supported by the Hungarian National Foundation for Scientific Research, Grant No. T 043758, and by the Slovenian-Hungarian Intergovernmental Scientific and Technological Cooperation Project, Grant No. SLO-1/03.

There are several results about sets which are contained in the union of three lines and have some other properties. For example Cameron [5] and Szőnyi [13] gave complete description of minimal blocking sets of this type.

The aim of this paper is to characterize the semiovals which are contained in the union of at most three lines. We will use the following notation throughout this note: Π is a projective plane of order q, S is a semioval in Π, if Q is a point of S then t_Q is the unique tangent to S at Q, $\mathcal{P}_Q$ is the pencil of lines with carrier Q, ℓ_1, ℓ_2 and ℓ_3 are the three lines whose union contains S, $L_i = S \cap \ell_i$ for $i = 1, 2, 3$, and $P_i = \ell_k \cap \ell_j$ where $\{i, j, k\} = \{1, 2, 3\}$.

2. Preliminaries

For $q = 2$ it is not hard to show that each semioval consists of three non-collinear points. Hence from now on we may assume that $q > 2$. It follows from the definition that a semioval could not be contained in one line. Suppose now that S is contained in the union of two lines, ℓ_1 and ℓ_2. Among the elements of $\mathcal{P}_{P_3}$ there exist $(q + 1) - 2 = q - 1$ lines which are tangent to S at P_3, so if $q > 2$ then $P_3 \notin S$. Let us choose an arbitrary point $Q \in L_1$. Then $q - 1$ out of the q lines of $\mathcal{P}_Q \setminus \ell_1$ must intersect ℓ_2 hence $|L_2| = q - 1$, and because of the symmetry $|L_1| = q - 1$. If $Q_i \in \ell_i$ are arbitrary points ($i = 1, 2$), then the pointset $\ell_1 \cup \ell_2 \setminus \{P, Q_1, Q_2\}$ is a semioval, because for each $R_i \in S$ the unique tangent t_{R_i} is the line $R_i Q_j$ where $\{i, j\} = \{1, 2\}$. Hence we proved the following:

PROPOSITION 2.1. *Let S be a semioval in a projective plane of order $q > 2$. If S is contained in the union of two lines ℓ_1 and ℓ_2, then $|S| = 2(q-1)$ and $S = \ell_1 \cup \ell_2 \setminus \{\ell_1 \cap \ell_2, Q_1, Q_2\}$ where $Q_i \in \ell_i$ for $i = 1, 2$.* ∎

If S is contained in the union of three lines, then there are much better bounds on the size of S than the general ones.

PROPOSITION 2.2. *Let S be a semioval in a projective plane Π of order q. If S is contained in the union of three lines then*

$$\frac{3(q-1)}{2} \leq |S| \leq 3(q-1).$$

PROOF. We may assume that $q > 4$ because if $q \leq 4$, then the bounds of Hubaut are sharper than the bounds of our proposition. The upper bound is a trivial consequence of a theorem of Dover [6]. He proved that if S is a semioval in a projective plane Π of order $q > 3$ and ℓ is any line of Π then $|S \cap \ell| \leq q - 1$.

In the case of the lower bound we distinguish two possibilities. If ℓ_1, ℓ_2 and ℓ_3 are concurrent, then their point of intersection $P_1 (= P_2 = P_3)$ does not belong to S, because $P_1 \in S$ would imply that there were $(q + 1) - - 3 > 2$ tangents to S at P_1. Let now $Q \in L_i$ be any point of S. Among the $q + 1$ lines of $\mathcal{P}_Q$ there are two exeptional ones, t_Q and ℓ_i, each of the remaining $q - 1$ lines meets either L_j or L_k where $\{i, j, k\} = \{1, 2, 3\}$. Thus $|L_j| + |L_k| \geq q - 1$. This holds for all the three possible pairs (j, k), hence $|L_1| + |L_2| + |L_3| \geq 3(q - 1)/2$.

If ℓ_1, ℓ_2 and ℓ_3 form a triangle, and $P_i \notin S$ then the same argument shows that $|L_j| + |L_k| \geq q - 1$. If $P_i \in S$ then $|L_i| \geq q - 2$, because among the lines of $\mathcal{P}_{P_i}$ there is only one, t_{P_i}, which does not contain some other points of S. Let $Q \in L_i$ be an arbitrary point. Now we get $|L_j| + |L_k| \geq q - 2$. Hence in both cases $|L_i| + |L_j| + |L_k| \geq 3(q - 1)/2$. ∎

In the rest of the paper semiovals in $PG(2, q)$ which are contained in the union of three lines are studied. We assume that S is not contained in the union of two lines, thus $L_i \setminus \{P_j, P_k\} \neq \emptyset$ for $\{i, j, k\} = \{1, 2, 3\}$. In Section 3 a complete classification is given when the lines form a triangle. We prove that each semioval belongs to one of the following three classes.

1. S has a $(q - 2)$-secant and two $(t + 1)$-secants for a suitable t. A semioval in this class exists if and only if $q = 4$ and $t = 1$, $q = 8$ and $t = 4$ or $q = 32$ and $t = 26$.
2. S has two $(q - 1)$-secants and a k-secant. Semiovals in this class exist for all $1 < k < q$.
3. S has three $(q - 1 - d)$-secants. Semiovals in this class exist if and only if $d | (q - 1)$.

In Section 4 some results are given when the lines are concurrent.

3. Semiovals contained in the sides of a triangle

We show that if ℓ_1, ℓ_2 and ℓ_3 form a triangle, then S belongs to one of classes 1–3 of semiovals on the list at the end of the previous section.

PROPOSITION 3.1. *S contains at most one point from the set $\{P_1, P_2, P_3\}$.*

PROOF. If $P_i \in S$ then $|L_i \setminus \{P_j, P_k\}| = q - 2$. Thus $\{P_1, P_2, P_3\} \subset S$ implies $|L_i| = q$, contradicting to the previously cited theorem of Dover. Suppose now that $P_1, P_2 \in S$ and $P_3 \notin S$. Then $|L_1| = |L_2| = q - 1$. Let E_i $(i = 1, 2)$ be the unique point of ℓ_i which is not in L_i and different from P_3. For each $A \in L_1$ t_A must be the line AE_2, hence $AE_2 \cap \ell_3 \notin S$, so L_3 contains exactly three points: P_1, P_2 and $E_1E_2 \cap \ell_3 = E_3$. But at E_3 there are two distinct tangents to S, the lines E_3P_3 and E_3E_1. This contradiction proves the statement. ∎

THEOREM 3.2. *A semioval in $PG(2, q)$ which is contained in the sides of a triangle and which contains one vertex of this triangle has a $(q-2)$-secant and two $(t+1)$-secants where t is a suitable integer. This type of semiovals exists if and only if $q = 4$ and $t = 1$, $q = 8$ and $t = 4$ or $q = 32$ and $t = 26$.*

PROOF. If S contains P_3 then Proposition 3.1 implies that neither P_1 nor P_2 are in S and $|L_3| = q - 2$. Hence there exists a point Q such that $\ell_3 \setminus L_3 = \{P_2, P_3, Q\}$. Let us choose the system of reference such that

$$P_1 = (1, 0, 0), \qquad P_2 = (0, 1, 0), \qquad P_3 = (0, 0, 1), \qquad Q = (1, 1, 0).$$

Let

$$A_1 = \{a \in GF^*(q) : (a, 0, 1) \in S\}$$

and

$$A_2 = \{a \in GF^*(q) : (0, -a, 1) \in S\}.$$

First we show that $A_1 = A_2$. If $R \in L_i$ is an arbitrary point $(i = 1, 2)$ then t_R is the line RP_i hence RQ contains at least two – and so exactly two – points of S. But the points $Q = (1, 1, 0)$, $(a, 0, 1)$ and $(0, -a, 1)$ are collinear. Thus $(a, 0, 1) \in S$ if and only if $(0, -a, 1) \in S$. Let now $t = |A_1| = |A_2|$.

If $1 \neq m \in GF^*(q)$ then $M = (m, 1, 0) \in L_3 \subset S$. Consider the elements of $\mathcal{P}_M$. The line ℓ_3 is a $(q-2)$-secant of S, t_M is a tangent, each of the remaining $q - 1$ lines is either a 2-secant or a 3-secant of S. Each 2-secant contains one point of $L_1 \cup L_2$ while each 3-secant contains one point of L_1 and one point of L_2. The cardinality of $L_1 \cup L_2$ is $2t + 1$, so if the number of 3-secants is λ, then $2\lambda + (q - 1 - \lambda) = 2t + 1$. Hence there are exactly $\lambda = 2t + 2 - q$ 3-secants of S in $\mathcal{P}_M$.

A 3-secant contains the points $(b, 0, 1)$, $(0, -c, 1)$ and $(m, 1, 0)$ if and only if $m = b/c$. Hence S is a semioval if and only if for all $1 \neq m \in GF^*(q)$ there exist exactly $\lambda = 2t + 2 - q$ pairs of elements (b, c) of $A_1 \times A_1$ for which $m = b/c$ hold. This means that A_1 is a difference set in $GF^*(q)$ with parameters $v = q-1, k = t, \lambda = 2t+2-q$. For the basic facts about difference sets we refer to the survey of Baumert [2].

If a (v,k,λ)-difference set exists, then its parameters satisfy the equation $k(k-1) = (v-1)\lambda$, hence in our case

$$t(t-1) = (q-2)(2t+2-q).$$

Solving this equation and using $t < q$ we get the parameters of the difference set:

$$v = q-1, \qquad k = q - \frac{3+\sqrt{4q-7}}{2}, \qquad \lambda = q-1-\sqrt{4q-7}.$$

Thus if $n = k - \lambda$ then

$$n^2+n+1 = \frac{4q-7-2\sqrt{4q-7}+1}{4} + \frac{\sqrt{4q-7}-1}{2} + 1 = q-1,$$

so the difference set is a planar one.

If q is odd then $4q-7 \equiv 5 \pmod 8$, hence $4q-7$ is not a square. Thus this type of difference set does not exist for q odd. So semiovals belonging to this class could exist only for q even. If q is even then $4q-7$ is a square if and only if $4q = 2^r$ and the diophantine equation $2^r = x^2+7$ has a solution. This equation was solved by Nagell [11]. He proved that there are five solutions, namely the pairs $(r,x) = (3,1)$, $(4,3)$, $(5,8)$, $(7,11)$, and $(15,181)$.

If $r = 3$ then $q = 2$, contrary to our assumption $q > 2$. If $r = 4$ then $q = 4$ and $\lambda = 0$, so there is no three-secant, the semioval contains five points, it is an oval. If $r = 5$ then $q = 8$ and the difference set has parameters $v = 7$, $k = 4$ and $\lambda = 2$. A difference set with these parameters exists, this is the complementary difference set of the well-known $(7,3,1)$-difference set belonging to the Fano plane. The corresponding semioval in $PG(2,8)$ consists of 15 points, it has two 5-secants and one 6-secant. If $r = 7$ then $q = 32$ and the difference set has parameters $v = 31$, $k = 25$ and $\lambda = 20$. Such difference set exists, this is the complementary difference set of the $(31,6,1)$-difference set which belongs to the projective plane of order $q = 5$. Hence the semioval appears in $PG(2,32)$. It has 81 points, two 26-secants and one 30-secant. If $r = 13$ then $q = 8192$ and the parameters are $v = 8191$, $k = 181$, $\lambda = 91$ and $n = 90$. There is no planar difference set with these parameters, because it is known (see [8]) that for $n < 2{,}000{,}000$ the order of each cyclic projective plane is a prime power. ∎

Now consider the cases when S does not contain any point from the set $\{P_1, P_2, P_3\}$. The vertexless triangle T is a semioval belonging to this class.

Let D be any set of points on one side of T. If $0 < |D| < q - 2$, then it is easy to show that the set $T \setminus D$ is a semioval. These semiovals form Class 2. If we delete points from more than one side of T, then the semioval belongs to Class 3.

THEOREM 3.3. *If a semioval S in $PG(2,q)$ is contained in the sides of a triangle T, does not contain any vertex of T and has at most one $(q-1)$-secant, then S has exactly three $(q-1-d)$-secants where d is a suitable divisor of $q-1$.*

PROOF. Let us choose the system of reference such that the lines ℓ_1 and ℓ_2 are not $(q-1)$-secants. Then we may assume that $P_1 = (1,0,0), P_2 = (0,1,0), P_3 = (0,0,1)$, and the points $(1,0,1)$ and $(0,1,1)$ are not in S. Let

$$A = \{a \in GF^*(q) : (a,0,1) \notin S\},$$
$$B = \{b \in GF^*(q) : (0,b,1) \notin S\}$$

and

$$C = \{c \in GF^*(q) : (-c,1,0) \notin S\}.$$

We prove that $A = B = C$. If $Q_i \in L_i$ then t_{Q_i} is the line $Q_i P_i$ for $i = 1,2,3$. Thus if two points, U and V from two distinct sides of T are not in S, W denotes the point of intersection of the line UV and the third side of T, then W could not be in S because the line UV would be another tangent through W. The points $(a,0,1)$, $(0,b,1)$ and $(-c,1,0)$ are collinear if and only if $a = bc$. Hence $a \in A$ and $b \in B$ imply $a/b \in C$, $a \in A$ and $c \in C$ imply $a/c \in B$, and $c \in C$ and $b \in B$ imply $bc \in A$. So $1 \in C$, because $1 \in A \cap B$. But this means that $A \subset B$ and $B \subset A$, hence $A = B$. In the same way we get $A = C$. Hence $a \in A$ and $b \in A$ imply $ab \in A$, and $1 \in A$ and $a \in A$ imply $1/a \in A$. This means that A is a subgroup of $GF^*(q)$.

If $G \neq GF^*(q)$ is an arbitrary subgroup, then the pointset

$$\{(h,0,1),(0,h,1),(-h,1,0) : h \in GF^*(q) \setminus G\}$$

is a semioval with cardinality $3(q-1-|G|)$, because the lines with equation $X_1 = hX_3$, $X_2 = hX_3$, $X_1 = -hX_2$ are the unique tangent lines at the points $(h,0,1)$, $(0,h,1)$, $(-h,1,0)$, respectively. ∎

4. Semiovals contained in three concurrent lines

If q is odd then the lower bound in Proposition 2.2 is sharp, there is a semioval with cardinality $3(q-1)/2$ in Class 3. If ℓ_1, ℓ_2 and ℓ_3 have a common point, P_1, then we can prove a slightly better lower bound on $|S|$.

THEOREM 4.1. *If a semioval S in $PG(2,q)$ is contained in the union of three concurrent lines then $|S| > 3(q-1)/2$ for $q > 9$.*

PROOF. If q is even then the statement follows from Proposition 2.2. Let q be odd and suppose that $|S| = 3(q-1)/2$. As we have already seen this implies $|L_1| = |L_2| = |L_3| = (q-1)/2$. So if $Q_i \in L_i$ for $i = 1,2,3$ then the points Q_1, Q_2 and Q_3 could not be collinear. Let us choose the system of reference such that the lines ℓ_1, ℓ_2 and ℓ_3 have equations $X_1 = -X_3$, $X_1 = 0$ and $X_1 = X_3$, respectively. Then $P_1 = (0,1,0) \notin S$. Let

$$A = \{a \in GF(q) :\ (-1,a,1) \in L_1\},$$
$$B = \{b \in GF(q) :\ (0,b,2) \in L_2\}$$

and

$$C = \{c \in GF(q) :\ (1,c,1) \in L_3\}.$$

Now we can consider the sets A, B and C as subsets of the additive group of $GF(q)$. We have $(A+C) \cap B = \emptyset$, otherwise $a+c=b$ would imply that the points $(-1,a,1), (0,b,2)$ and $(1,c,1)$ were collinear. Hence $|A+C| \leq (q+1)/2$. But the Theorem of Kneser (see [10], p. 6.) states that there exists a subgroup H such that $A+C = A+C+H$ and $|A+C| \geq |A+H| + |C+H| - |H|$. So $(q+1)/2 \geq |H| \geq (q-3)/2$. The order of a subgroup divides the order of the group, so $1 < |H|$ divides q. But $2|H| \neq q$ and $3(q-3)/2 > q$ if $q > 9$, so there is no such semioval for $q > 9$. ∎

It is easy to see that Theorem 4.1 is valid for $q = 3, 7$ and 9, too. For $q = 5$ each oval contains $q + 1 = 3(q-1)/2 = 6$ points. If P_1 is an internal point of an oval, then the oval is contained in the three secants passing on P_1, so this is the only case when Theorem 4.1 is not true.

We were able to construct only one infinite class of this type of semiovals. This is the following.

EXAMPLE 4.2. Let $q = s^2$ and let ℓ_1, ℓ_2, ℓ_3 be three concurrent lines in $PG(2,q)$. For $i = 1,2,3$ choose $\overline{\ell}_i \subset \ell_i$ Baer sublines such that each Baer subplane $\langle\overline{\ell_i},\overline{\ell_j}\rangle$ meets the line ℓ_k only in P_1 if $\{i,j,k\}=\{1,2,3\}$. Then $S=(\ell_1\setminus\overline{\ell}_1)\cup(\ell_2\setminus\overline{\ell}_2)\cup(\ell_3\setminus\overline{\ell}_3)$ is a semioval which has $3(q-\sqrt{q})$ points.

The line ℓ_i is tangent to the Baer subplane $\mathcal{B}_{j,k}=\langle\overline{\ell_j},\overline{\ell_k}\rangle$ if $\{i,j,k\}=$ $=\{1,2,3\}$. Hence $s+1$ lines of $\mathcal{B}_{j,k}$ pass on P_1 and exactly one line of $\mathcal{B}_{j,k}$ passes on each other point of ℓ_i. So for each point $Q\in L_i$ there is a unique line of $\mathcal{B}_{j,k}$ which passes on Q. This line is t_Q, because any other element of $\mathcal{P}_Q$ does not belong to the set of lines of $\mathcal{B}_{j,k}$, thus it meets $(\ell_j\setminus\overline{\ell_j})\cup(\ell_k\setminus\overline{\ell_k})=L_j\cup L_k$ in at least one point.

We can construct such a semioval for example in the following way. Let i be a root of an irreducible quadratic polynomial of $GF(s)[X]$ and consider $GF(q)$ as the extension of $GF(s)$ by i. The equations of the lines are as follows: $\ell_1: X_2=0$, $\ell_2: X_1=0$ and $\ell_3: X_1=iX_2$, and the Baer sublines are:

$$\begin{aligned}\overline{\ell}_1&=\{(a,0,1):a\in GF(s)\}\cup\{(1,0,0)\},\\ \overline{\ell}_2&=\{(0,b,1):b\in GF(s)\}\cup\{(0,1,0)\},\\ \overline{\ell}_3&=\{(1,i,ci+1):c\in GF(s)\}\cup\{(0,0,1)\}.\end{aligned}$$

If $s=2$ then $q=4$ and our example has $6=q+2$ points. Semiovals with cardinality $q+2$ were studied by Blokhuis [3]. He proved that these objects exist if and only if $q=4$ or 7. If $q=7$, then there is a projectively unique semioval which contains nine points. This semioval is contained in the union of three non-concurrent lines and belongs to Class 3 on the list at the end of Section 2.

References

[1] L. M. BATTEN, Determining sets, *Australas. J. Combin.,* **22** (2000), 167–176.

[2] L. D. BAUMERT, Cyclic difference sets, *Lecture notes in Mathematics,* **182**, Springer, Berlin–Heidelberg–New York, 1971.

[3] A. BLOKHUIS, Characterization of seminuclear sets in a finite projective plane, *J. Geom.,* **40** (1991), 15–19.

[4] A. BLOKHUIS and T. SZŐNYI, Note on the structure of semiovals in finite projective planes, *Discrete Math.,* **106/107** (1992), 61–65.

[5] P. J. CAMERON, Four lectures on Projective Geometries, in: Finite Geom. (ed.: Baker, C.A. and Batten, L.M.) *Lecture Notes in pure and applied math.,* **103**, Marcel Dekker (1985), 27–63.

[6] J. M. DOVER, Semiovals with large collinear subsets, *J. Geom.,* **69** (2000), 58–67.

[7] A. GÁCS, On regular semiovals, *J. Algebraic Combin.,* submitted.

[8] D. M. GORDON, The prime power conjecture is true for $n < 2,000,000$, *Electronic J. Combin.,* **1 R6** (1994).

[9] X. HUBAUT, Limitation du nombre de points d'un (k,n)-arc régulier d'un plan projectif fini, *Atti Accad. Naz. Lincei Rend.,* **8** (1970), 490–493.

[10] H. B. MANN, *Addition Theorems,* John Wiley, 1965.

[11] T. NAGELL, The diophantine equation $x^2+7=2^n$, *Ark. Mat.,* **4** (1961), 185–187.

[12] C. SUETAKE, Two families of blocking semiovals, *European J. Combin.,* **21** (2000), 973–980.

[13] T. SZŐNYI, Combinatorial problems for abelian groups arising from geometry, *Period. Polytech. Transportation Engrg.,* **19** (1991), no. 1-2, 91–100.

[14] J. A. THAS, On semiovals and semiovoids, *Geom. Dedicata,* **3** (1974), 229–231.

György Kiss

Department of Geometry,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
and
Bolyai Institute, University of Szeged
H–6720 Szeged, Aradi vértanúk tere 1,
Hungary
kissgy@cs.elte.hu

János Ruff

Bolyai Institute, University of Szeged
H–6720 Szeged, Aradi vértanúk tere 1,
Hungary
ruffj@math.u-szeged.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 107–113

NOTES ON THE EXISTENCE OF INVARIANT MORSE FUNCTIONS

By

B. DOBAI

(Received 26 October, 2004)

Introduction

In the definition of Morse functions the requirement that the critical points are non-degenerate, can be replaced by the more general one that the critical points form submanifolds which are non-degenerate in a sense defined below [1]. Thus a concept of generalized Morse function is obtained. The problem concerning to the existence of invariant Morse functions was fully resolved by Wasserman [11], by means of abstract existence theorems without constructing any example. Riemannian manifolds with isometric actions which admit orthogonally transverse submanifolds are applied below to construct examples of invariant generalized Morse functions.

1. Some basic facts

DEFINITIONS. A connected submanifold $K \subseteq M$ is a *critical submanifold* of the smooth function $f : M \longrightarrow \mathbb{R}$ if every point z of K is a critical point of f.

Let $K \subseteq M$ be a critical submanifold then we have the following inclusion:

(1) $T_z K \subseteq \{v \in T_z M | Hessf(v, T_z M) = \{0\}\}$ for every $z \in K$.

Furthermore if in (1) instead of inclusion equality holds for every $z \in K$ then the considered submanifold is called *non-degenerate critical submanifold* of f.

Let G be a Lie group and $\phi : G \times M \longrightarrow M$ be a smooth action of G on a smooth manifold M. A smooth function $f : M \longrightarrow \mathbb{R}$ is said to be *G-invariant smooth function* if:

$$f(\phi(g,x)) = f(x)$$

for every $x \in M$ and $g \in G$.

If $x \in M$ is a critical point of a G-invariant smooth function, where G is a compact Lie group, the orbit of x is itself a critical submanifold of f, called *critical orbit*.

Let K be a critical submanifold of M if the inclusion $\phi(G, K) \subseteq K$ holds then K is called an *invariant critical submanifold* of M [1].

We apply the following well-known result [2]:

PROPOSITION 1. *Let ϕ be a smooth action of a compact Lie group G on a smooth manifold M, then for every $x \in M$ we can construct a slice, i.e. a submanifold S_x of M with the following properties:*

(1) The saturation of S_x i.e. the set $\{\phi(g,z) \mid z \in S_x , g \in G\}$ is an open tubular neighbourhood of $G(x)$, furthermore $S_x \cap G(x) = \{x\}$ is fulfilled;
(2) For every $g \in G_x$ we have $\phi(g, S_x) = S_x$;
(3) If $z \in S_x$ then $G_z \subseteq G_x$;
(4) If $\phi(g, S_x) \cap S_x$ is non-empty then g belongs to G_x.

We also apply the construction of normal slices which can be done as follows: Choose a G-invariant Riemannian metric $\langle\ ,\ \rangle : TM \times TM \longrightarrow \mathbb{R}$ on M, such a metric exits as a consequence of the existence of an invariant Haar measure on the group G. We can decompose the tangent space $T_z M$ at $z \in G(x)$ into two orthogonally complementary subspaces: $T_z G(x)$, respectively $T_z^{\perp} G(x)$, the last one will be denoted throughout this paper by $\nu_z G(x)$.

Let $\nu_z^r G(x)$ be the subset of $\nu_z G(x)$ whose elements satisfy the condition $\|v\|_z < r$ where r is a positive number. Then by a suitable choice of $r > 0$ we can produce the slice S_z as the image of $\nu_z^r G(x)$ under the exponential map. For every slice S_z which has been constructed as above we have

$$T_z S(z) = \nu_z G(x).$$

Since every vector $v \in \nu_z G(x)$ can be thought as a tangent vector to a segment of geodesic lying in S_z, the inclusion follows: $\nu_z G(x) \leq T_z S_z$; moreover S_z has obviously the dimension $dim\, M - dim\, G(x)$, therefore we have $T_z S_z = \nu_z(x)$. Finally if the vector $v \in T_z M$ belongs to $T_z S_z$ and $\gamma : [0, 1] \longrightarrow M$ is a geodesic with $\dot{\gamma}(0) = v$ then it has a piece $\gamma\left([0, \delta]\right)$ with $\delta > 0$ included in S_z, thus the submanifold S_z is geodesic at the point x.

COROLLARY 1. *Every normal slice of an orbit* $G(x)$ *is geodesic in* x.

2. The main result

PROPOSITION 2. *Let* ϕ *be a smooth action of a compact Lie group* G *on a smooth manifold* M *furthermore let* $f : M \longrightarrow \mathbb{R}$ *be a* G*-invariant smooth function. Consider now a critical orbit* $G(x)$ *with respect to* f *and let* S_x *be the normal slice of the orbit* $G(x)$. *Then* $G(x)$ *is a non-degenerate critical submanifold of* f *if and only if* x *is a non-degenerate critical point of the restricted function* $f \restriction_{S_x}$.

This propositon is a particular case of the following more general result ($theorem$, 1).

DEFINITION 1. In case of an isometric action a submanifold $L \subseteq M$ is said to be an *orthogonally transverse submanifold* to the action ϕ if the following two conditions are satisfied:

(1) The submanifold L intersects every orbit $G(z)$, $z \in M$ of the action;

(2) The subspaces $T_z L$, $T_z G(z) \subseteq T_z M$ are orthogonal to each other at every point $z \in L$.

There are given fairly general conditions which assure the existence of orthogonlly transverse submanifolds. For details see [8].

The statement of the first proposition can be extended in the case when instead of normal slice we have orthogonally transverse submanifold.

THEOREM 1. *Let* $\phi : G \times M \longrightarrow M$ *be an isometric action which admits an orthogonally transverse submanifold* L *and* $f : M \longrightarrow M$ *an invariant smooth function. Then the critical orbits of* ϕ *are non-degenerate submanifold if and only if* $f \restriction L$ *is a Morse function.*

The proof can be made step by step through the following lemmas:

LEMMA 1. *Let* $L \subseteq M$ *be an orthogonally transverse submanifold of the action* ϕ *and* $f : M \longrightarrow \mathbb{R}$ *a smooth invariant function. Then the gradient field of* f *is a smooth extention of the gradient of* $f \restriction_L$.

PROOF. Since f is invariant, $\operatorname{grad} f \in \nu_z G(z)$. If $z \in L$ is such that $G(z)$ is principal orbit, then $\nu_z G(z) = T_z L$. But $\operatorname{grad}(f \restriction_L)$ is the orthogonal projrection of $\operatorname{grad} f$ on TL, and taking into account the principal orbit type theorem we have $\operatorname{grad}(f \restriction_L) = (\operatorname{grad} f) \restriction_L$. ∎

LEMMA 2. *By keeping the assumptions made above for the gradient of the function* $f\restriction_L$ *we have:*

$$T_z\phi_g(\mathrm{grad}(f\restriction_L)) = \mathrm{grad}(f\restriction_{g(L)}(\phi_g(z)))$$

where $g(L) = \phi(g, L)$

PROOF. Clearly for every $g \in G$ we have $T_z\phi_g(T_z L) = T_{\phi_g(z)}(g(L))$. So we can write

$$T_z\phi_g(v)(f) = v(f \circ \phi_g) = \langle \mathrm{grad}\, f \circ \phi_g, v\rangle = \langle T_z\phi_g(v), \mathrm{grad}(f(\phi_g(z)))\rangle$$

for every tangent vector $v \in T_z L$.

On the other hand ϕ_g for $g \in G$ is an isometry of M therefore

$$\langle T_z\phi_g(v), T_z\phi_g(\mathrm{grad}(f))\rangle = \langle v, \mathrm{grad}(f)\rangle$$

thus

$$\langle T_z\phi_g(v), T_z\phi_g(\mathrm{grad}(f)) - \mathrm{grad}(f(\phi_g(z))\rangle = 0;$$

this last equation and the first lemma yield that

$$T_z\phi_g(\mathrm{grad}(f\restriction_L)) = \mathrm{grad}(f\restriction_{g(L)}(\phi_g(z))). \qquad \blacksquare$$

LEMMA 3. *The restricted function* $f\restriction_L\colon L \longrightarrow \mathbb{R}$ *has only non-degenerate critical points if and only if* $f\restriction_{g(L)}\colon g(L) \longrightarrow \mathbb{R}$ *also has only non-degenerate critial points.*

PROOF. The tangend space $T_z L$ can be decomposed into directsum $T_z L = \bigoplus_\lambda R_\lambda$ of the eigenspaces of the Hessian of $f\restriction_L$. Clearly the linear operator which belongs to the bilinear form $\mathrm{Hess}\, f\restriction_L(z)$ can be written as $\nabla_{(.)}\mathrm{grad}(f\restriction_L)$. By choosing an eigenvector $v \in R_\lambda$ we can write

$$\begin{aligned}
\nabla_{T_z\phi_g(v)}\mathrm{grad}(f\restriction_{g(L)}) &= \nabla_{T_z\phi_g(v)} T_z\phi_g(\mathrm{grad}(f\restriction_L) \\
&= T_z\phi_g \circ (T_z\phi_g^{-1} \circ \nabla_{T_z\phi_g(v)} T_z\phi_g)(\mathrm{grad}(f\restriction_L)) \\
&= T_z\phi_g \circ ((T_z\phi_g)^*\nabla)_v\, \mathrm{grad}(f\restriction_L) \\
&= T_z\phi_g \circ \nabla_v\, \mathrm{grad}(f\restriction_L) = T_z\phi_g(\lambda v) = \lambda T_z\phi_g(v),
\end{aligned}$$

therefore we have $T_z\phi_g(R_\lambda) \subseteq R_\lambda'$ where R_λ' denotes the corresponding eigensubspace of $\mathrm{Hess}\, f\restriction_{g(L)}(z)$, which proves the assertion.

PROOF OF THE THEOREM. For each $u, v \in T_z M$ we have

$$\mathrm{Hess}\, f(z)(u, v) = \langle \nabla_u\, \mathrm{grad}\, f; v\rangle = \langle \alpha(u); v\rangle,$$

here $\alpha : T_z M \longrightarrow T_z M$ and below $\beta : T_z L \longrightarrow T_z L$ denote linear operators belonging to the Hessian of f resp. $f \restriction_L$.

Now we claim that $\alpha \restriction_{T_z L} = \beta$.

Indeed since L is totally geodesic with its induced covariant derivation $\tilde{\nabla}$ and second fundamental form ω the following holds:

$$\begin{aligned}\beta(u) &= \tilde{\nabla}_u \operatorname{grad}(f \restriction_L) = \bar{\nabla}_u \operatorname{grad}(f \restriction_L) - \omega(u, \operatorname{grad}(f \restriction_L)) \\ &= \bar{\nabla}_u \operatorname{grad}(f \restriction_L)\end{aligned}$$

where $\bar{\nabla}$ is the restriction of ∇ to L.

By taking into account the lemma 1 from above we can write:

$$\beta(u) = \bar{\nabla}_u \operatorname{grad}(f \restriction_L) = \nabla_u(\operatorname{grad} f) \restriction_L = \alpha(u).$$

On the other hand by a result of Conlon [3], for each $v \in \nu_z G(z)$ there is $g \in G_z$ such that $v \in T_z\phi_g(T_z L) = T_z g(L)$ thus α is not singular on $T_z g(L)$ that is

$$\operatorname{Hess} f(z) \restriction_{\nu_z G(z) \times \nu_z G(z)}$$

is non-degenerate.

Conversely supposing that $G(z)$ is non-degenerate orbit and furthermore z is a degenerate critical point of $f \restriction_L$ the we have a nonzero vector $u \in T_z L$ such that $\beta(u) = 0$ that is $\alpha(u) = 0$ however $T_z L \leq \nu_z G(z)$ in contradiction with the assumption that $G(z)$ is non-degenerate.

3. Some examples

Let L be an orthogonally transverse submanifold of a smooth manifold M then there exists a finite group W, called the *generalized Weyl group* of S which acts on L so that there is a bijection $L/W \backsimeq M/G$ between the orbit spaces realized by $\Psi : W(x) \mapsto G(x)$ for each $x \in S$. For details see [8], [6].

The problem to find G invariant smooth functions on M can be tranposed to the simpler one to seek W invariant smooth functions on L. In order to resolve this we need the following result due to Palais and Terng ([6]).

THEOREM 2. *Let L be an orthogonally transverse submanifold for the Riemannian G-manifold M and let W be its generalized Weyl group. Then the restriction map $f \mapsto f \restriction_L$ is an isomorphism between the Banach algebras*

$$C^\infty(M)^G \longrightarrow C^\infty(S)^W,$$

i.e. between the Banach algebras of G resp. W invariants smooth functions on M resp. L.

As a corollary of the above theorem every W invariant smooth function $f : L \longrightarrow \mathbb{R}$ has an unique smooth, G-invariant extention $\widetilde{f} : M \longrightarrow \mathbb{R}$.

Here the smoothness of such a extension requires some deeper considerations see [6].

Now we are able to construct in some special cases non-degenerate invariant Morse functions.

EXAMPLE. Consider now a compact, connected, semi-simple Lie group G and its Lie algebra $\mathfrak{g}$. Then G acts on $\mathfrak{g}$ by

$$\phi(g;X) = T_e ad(g)X, \quad \text{where} \quad g \in G \quad \text{and} \quad X \in \mathfrak{g}\,.$$

This is an orthogonal action with respect to the Cartan–Killing form furthermore it has an orthogonally transverse submanifold $\mathfrak{l}$ which actually is a Cartan subalgebra of $\mathfrak{g}$ (see e.g. [10]).

As $\mathfrak{l}$ is an euclidean space the corresponding Weyl group W acts on $\mathfrak{l}$ as a Coxeter group generated by reflections. In the case when W is isomorphic to A_n the symmetric group, it acts by permuting $x_1, x_2, x_3, \dots, x_{n+1}$ subject to the relation $x_{n+1} = -(x_1 + \cdots + x_n)$. We let

$$f_i := x_1^{i+1} + \cdots + x_{n+1}^{i+1} \quad (1 \leq i \leq n)$$

these polinomials are invariant respect to W (see e.g. [5]). It is easy to see that numbers α and β can be chosen so that $\alpha f_j + \beta f_k$, $l \neq k$ has only nondegenerate critical points.

DEFINITION 2. The invariant smooth function $f : M \longrightarrow \mathbb{R}$ is called *Morse function* for the Riemannian G-manifold if the critical locus of f is a union of non-degenerate critical manifolds without interior.

Since the action considered above is not transitive, it has not orbits with non-empty interior therefore we have an invariant Morse function in sense of Wassermann [11].

References

[1] R. BOTT, Non-degenerate Critical Manifolds, *Ann. Math.* **60** (1954), 248–261.

[2] G. E. BREDON, *Introduction to Compact Transformation groups*, New York, 1972.

[3] L. CONLON, Variational completness and K-transversal domains, *J. Differential Geometry* **5** (1971), 135–147.

[4] R. Hermann, On the existence of a fundamental domain for Riemannian transformation groups, *Proc. Amer. Math. Soc.* **13** (1962), 489–494.

[5] James E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Univ. Press, 1990.

[6] Richard S. Palais and Chuu-Lian Terng, A general theory of canonnical forms, *Trans. Amer. Math. Soc.* **300** (1987), 771–789.

[7] Richard S. Palais and Chuu-Lian Terng, *Critical Point Theory and Submanifold Geometry*, Berlin, 1988.

[8] J. Szenthe, Orthogonally transversal submanifolds and the generalization of the Weyl group, *Period. Math. Hungar.* **50** (1984), 281–299.

[9] J. Szenthe, Some isometric actions with orthogonally transversal submanifolds on Riemannian symmetric spaces, *Stud. Scien. Math. Hungar.* **21** (1986), 175–179.

[10] J. Szenthe, Sur les voisinages des orbites d‘actions orthogonales, *Publ. Math.* **29** (1982), 43–55.

[11] Arthur G. Wasserman, Equivariant differential topology, *Topology* **8** (1969), 127–150.

B. Dobai

Demartment of Geometry,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
dobarna@cs.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 115–121

A DIVISIBILITY PROBLEM OF BINOMIAL COEFFICIENTS

By

GÁBOR NYUL

(Received October 26, 2004)

1. Introduction

Several results have been proved about binomial coefficients. We mention two of them, a diophantine problem and one about prime factors of binomial coefficients.

It was an interesting question when a binomial coefficient $\binom{n}{k}$ is a perfect power. Apart from the trivial cases $k \in \{0, 1, n-1, n\}$ it was proved that for $k = 2$ and $k = n-2$ the binomial coefficient $\binom{n}{k}$ is a square-number for infinitely many n. In the remaining cases the only solutions are $\binom{50}{3} = \binom{50}{47} = 140^2$ (for a survey see e.g. [3]).

J. J. Sylvester [5] and I. Schur [4] independently showed that if $n \geq 2k$, then the greatest prime divisor of $\binom{n}{k}$ is greater than k. As a generalization E.F.Ecklund Jr., R.B.Eggleton, P.Erdős and J.L.Selfridge [1] proved that if $n \geq 2k$, then the product of all prime factors $< k$ of $\binom{n}{k}$ is less than the product of all prime factors $\geq k$ of $\binom{n}{k}$ except 12 binomial coefficients.

In this paper we investigate the divisibility problem

$$nk \,\Big|\, \binom{n}{k} \quad (n, k \in \mathbb{Z},\ 0 \leq n,\ 0 \leq k \leq n) \tag{1}$$

of binomial coefficients. A more difficult problem is to solve the diophantine equation

$$\binom{n}{k} = bnk \quad (n, k, b \in \mathbb{Z},\ 0 \leq n,\ 0 \leq k \leq n,\ 1 \leq b). \tag{2}$$

Mathematics Subject Classification (2000): 11B65

In Section 2 we solve equation (2). First we give all solutions of our diophantine equation and the related inequality in (n,k) for $b=1$. If $b \neq 1$ then we give an upper bound for the solutions in (n,k). Finally we completely solve the equation in (n,b) for $k \leq 3$.

In Section 3 we prove that if $k=p$ is a prime number then the solutions of (1) in n belong to $p-1$ residue classes modulo p^2 and we show that a similar assertion is true for $k=4$.

2. Solving the diophantine equation $\binom{n}{k}=bnk$

2.1. Solutions in (n,k) with fixed b: $b=1$

THEOREM 1. *For $b=1$ the set of all solutions of (2) in (n,k) is*

$$\{(n,1) \mid 1 \leq n\} \cup \{(5,2)\}.$$

PROOF. For $k=0,1,2$ we can directly check that the solutions are $(n,1)$ with arbitrary $n \in \mathbb{N}$ and $(5,2)$.

Now it is enough to prove that there are no solutions with $k \geq 3$. Equation (2) is equivalent to

$$(n-1)\cdot \ldots \cdot (n-k+1) = k!\cdot k.$$

By $k>1$ we have

$$3\cdot \ldots \cdot (k+1) = \frac{(k+1)!}{2!} < k!\cdot k$$

and by $k>2$ we have

$$4\cdot \ldots \cdot (k+2) = \frac{(k+2)!}{3!} > k!\cdot k.$$

These inequalities imply that $k!\cdot k$ cannot be the product of $k-1$ consecutive integer, hence our equation has no solutions when $k \geq 3$. ∎

To have a complete description of the relation between $\binom{n}{k}$ and nk we prove the following theorem.

THEOREM 2. *The set of all solutions of*

$$\binom{n}{k} > nk \quad (n,k \in \mathbb{Z},\ 0 \leq n,\ 0 \leq k \leq n)$$

in (n,k) *is*

$$\{(n,0) \mid 0 \le n\} \cup \{(n,k) \mid 6 \le n \text{ and } 2 \le k \le n-3\}.$$

PROOF. For $k = 0, 1, 2, n-2, n-1, n$ we can easily verify that all solutions are $(n,0)$ with arbitrary $n \in \mathbb{N} \cup \{0\}$ and $(n,2)$ with arbitrary $n \ge 6$.

Now suppose that $3 \le k \le n-3$ (for the existence of such k let $n \ge 6$). Our inequality is equivalent to

$$(n-1) \cdot \ldots \cdot (n-k+1) > k! \cdot k.$$

By the proof of Theorem 1, case $k \ge 3$ and using $k \le n-3$ we have

$$k! \cdot k < 4 \cdot \ldots \cdot (k+2) \le (n-k+1) \cdot \ldots \cdot (n-1)$$

which means that our inequality holds for $n \ge 6$ and $3 \le k \le n-3$. ∎

2.2. Solutions in (n,k) with fixed b: $b \neq 1$

THEOREM 3. *If* $b \neq 1$ *is fixed, then (2) has finitely many solutions in* (n,k)*.*

Moreover $(n,k) = (4b+1, 2)$ *is a solution for arbitrary* b *and for all other solutions* (n,k) *we have* $6 \le n \le 6b$ *and* $3 \le k \le n-3$*.*

PROOF. When $b \neq 1$ is fixed then we can easily check that there are no solutions with $k = 0, 1, n-2, n-1, n$. For $k = 2$ the pair $(n,2)$ is a solution if and only if $n = 4b+1$.

Let $3 \le k \le n-3$ (and $n \ge 6$). Then

$$\binom{n}{3} \le \binom{n}{k} = bnk \le bn(n-3)$$

which is equivalent to $n^2 + (-3-6b)n + (18b+2) \le 0$. This implies that there are only finitely many suitable n and

$$n \le \frac{3 + 6b + \sqrt{36b^2 - 36b + 1}}{2} < \frac{3 + 6b + \sqrt{36b^2 - 36b + 9}}{2} = 6b. \qquad ∎$$

2.3. Solutions in (n,b) with fixed k: $k \leq 3$

THEOREM 4. *For $k = 0$ there are no solutions of (2) in (n,b). For $k = 1$ the set of all solutions of (2) in (n,b) is $\{(n,1) \mid 1 \leq n\}$. For $k = 2$ the set of all solutions of (2) in (n,b) is*

$$\{(n,b) \mid 1 \leq b, n = 4b + 1\}.$$

For $k = 3$ the set of all solutions of (2) in (n,b) is

$$\{(18t+2, 18t^2+t) \mid 1 \leq t\} \cup \{(18t+1, 18t^2-t) \mid 1 \leq t\} \cup$$
$$\{(18t+10, 18t^2+17t+4) \mid 0 \leq t\} \cup \{(18t-7, 18t^2-17t+4) \mid 1 \leq t\}.$$

PROOF. For $k = 0$ and $k = 1$ our statement is obvious. For $k = 2$ see the proof of Theorem 3.

For $k = 3$ equation (2) is equivalent to $n^2 - 3n + (2 - 18b) = 0$. It has integer root in n if and only if its discriminant, $72b + 1$ is a square-number, say m^2 ($m \in \mathbb{Z}$).

By the chinese remainder theorem the following congruences are equivalent: $m^2 \equiv 1 \pmod{72} \iff m^2 \equiv 1 \pmod{8}$ and $m^2 \equiv 1 \pmod{9} \iff m \equiv 1 \pmod{2}$ and $m \equiv \pm 1 \pmod{9} \iff m \equiv \pm 1 \pmod{18}$.

This implies $m = 36t \pm 1$ or $m = 36t \pm 17$ ($t \in \mathbb{Z}$), from which we get $b = 18t^2 \pm t, 18t^2 \pm 17t + 4$ and $n = 18t+2, 18t+1, 18t+10, 18t-7$ respectively. We have to notice that t must be non-negative and $t = 0$ is suitable only in the third case. ∎

3. Solving the divisibility problem $nk \mid \binom{n}{k}$

3.1. Solutions in n with fixed k: $k = p$ is a prime

THEOREM 5. *Let $k = p$ be a prime number and denote by $\overline{m}$ the residue class of m modulo p^2. Then the solutions of (1) in n are exactly those elements of $\overline{1} \cup \ldots \cup \overline{p-1}$ which are $\geq p$.*

We need two lemmas. They can be found in [2] with proofs (p. 505 and 507).

LEMMA 1. *If* $n, k \in \mathbb{Z}$, $0 \le n$, $0 \le k \le n$ *and* $\gcd(n,k) = 1$ *then* $n \mid \binom{n}{k}$.

LEMMA 2. *If* $n, p \in \mathbb{Z}$, $0 \le n$, $0 \le p \le n$ *and* p *is a prime number then* $\binom{n}{p} \equiv \left[\frac{n}{p}\right] \pmod{p}$.

PROOF OF THEOREM 5. Our purpose is to solve (1) in n. We distinguish two cases.

Case 1: $\gcd(n,p) = 1$. By this assumption (1) is equivalent to $n \mid \binom{n}{p}$ and $p \mid \binom{n}{p}$. The first divisibility holds by Lemma 1. The second divisibility is equivalent to $p \mid \left[\frac{n}{p}\right]$ by Lemma 2. This implies

$$n \equiv 1 \text{ or } \dots \text{ or } p-1 \pmod{p^2}.$$

Case 2: $\gcd(n,p) \ne 1$. We prove by induction that this implies $p^{2m} \mid n$ for every $m \in \mathbb{N}$ which is impossible.

By the assumption of this case $p \mid n$, hence $n = px$ $(x \in \mathbb{Z})$. By (1) $p \mid p^2x \mid \binom{px}{p}$ so Lemma 2 implies $p \mid \left[\frac{px}{p}\right] = x$, hence $p^2 \mid n$ which is our assertion for $m = 1$.

Suppose that $p^{2m} \mid n$ for $m \in \mathbb{N}$, that is $n = p^{2m}y$ $(y \in \mathbb{Z})$. By (1) we get

$$p^{2m+1} \mid p^{2m+1}y \mid \binom{p^{2m}y}{p} = \frac{p^{2m-1}y \cdot (p^{2m}y - 1) \cdot \ldots \cdot (p^{2m}y - p + 1)}{(p-1)!}.$$

It is possible only if $p^2 \mid y$, whence $p^{2m+2} \mid n$ is proved. ∎

3.2. Solutions in n with fixed k: arbitrary k

Theorem 5 suggests the following question. Is a similar assertion true for composite k? More precisely:

QUESTION. Do there exist $s, l \in \mathbb{N}$ and $\overline{m_1}, \dots, \overline{m_s}$ distinct residue classes modulo l for all $k \in \mathbb{N}$ such that the solutions of (1) in n are exactly those elements of $\overline{m_1} \cup \ldots \cup \overline{m_s}$ which are $\ge k$? (If possible, give $s, l, m_1, \dots, m_s$ explicitely.)

REMARK. For $k = p$ we proved in Theorem 5 that $s = p - 1$, $l = p^2$ and $m_i = i$ $(i = 1, \dots, s)$.

For prime-powers our conjecture is:

CONJECTURE. *For $k = p^{\alpha}$ (p is a prime, $\alpha \in \mathbb{N}$)*

$$s = \frac{p^{2\alpha} - 1}{p+1}, \quad l = p^{3\alpha - 1}.$$

Moreover let $A = \{t \in \mathbb{Z} \mid 1 \leq t \leq p^{\alpha} - 1\}$, $A' = \{a \in A \mid p \nmid a\}$ and $\mathcal{A} = \{\overline{a} \mid a \in A\} \cup \{\overline{a + j \cdot p^{2\alpha}} \mid a \in A'; j = 1, 2, \ldots, p^{\alpha - 1} - 1\}$ be a set of residue classes modulo $l = p^{3\alpha - 1}$. If $\mathcal{M} = \{\overline{m_1}, \ldots, \overline{m_s}\}$ is the set of residue classes modulo l containing the solutions of (1) in n for $k = p^{\alpha}$, then our additional conjecture is $\mathcal{A} \subseteq \mathcal{M}$.

REMARK. If the conjecture with its additional part is true, then we would have $\mathcal{A} = \mathcal{M}$ for $\alpha = 2$ (since $|\mathcal{A}| = p^{2\alpha-1} - p^{2\alpha-2} + p^{\alpha-1} - 1$ is equal to $s = \frac{p^{2\alpha}-1}{p+1}$ for $\alpha = 2$).

This conjecture is based on computational results. These suggest that the residue classes $\overline{m_1}, \ldots, \overline{m_s}$ for prime-powers $k < 10$ are

k	s	l	residue classes
4	5	32	$\overline{1}, \overline{2}, \overline{3}, \overline{17}, \overline{19}$
8	21	256	$\overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}, \overline{65}, \overline{67}, \overline{69}, \overline{71}, \overline{129}, \overline{130}, \overline{131}, \overline{133}, \overline{134}, \overline{135}, \overline{193}, \overline{195}, \overline{197}, \overline{199}$
9	20	243	$\overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}, \overline{8}, \overline{82}, \overline{83}, \overline{85}, \overline{86}, \overline{88}, \overline{89}, \overline{163}, \overline{164}, \overline{166}, \overline{167}, \overline{169}, \overline{170}$

Finally we prove our conjecture in the very special case $p = \alpha = 2$.

THEOREM 6. *Let $k = 4$ and denote by $\overline{m}$ the residue class of m modulo 32. Then the solutions of (1) in n are exactly those elements of $\overline{1} \cup \overline{2} \cup \overline{3} \cup \overline{17} \cup \overline{19}$ which are ≥ 4.*

PROOF. If we prove that $4n \mid \binom{n}{4}$ implies $4(n+32) \mid \binom{n+32}{4}$ for $n \in \mathbb{Z}$, $n \geq 4$, then obviously $l = 32$ is a good choice with the notation of our question.

The assumption $4n \mid \binom{n}{4}$ is equivalent to $96 \mid (n-1)(n-2)(n-3)$. But the product of three consecutive integer is always divisible by 3, whence we get $32 \mid (n-1)(n-2)(n-3)$ which implies $32 \mid (n+31)(n+30)(n+29)$. It is equivalent to our assertion that can be verified by the same steps.

Now it only remains to check which elements of $\{n \in \mathbb{Z} \mid 4 \leq n \leq 35\}$ are solutions of (1) in n. It turns out that there are five solutions in this set, namely 17, 19, 33, 34, 35. ∎

References

[1] E. F. ECKLUND JR., R. B. EGGLETON, P. ERDŐS and J. L. SELFRIDGE, On the prime factorization of binomial coefficients, *J. Austral. Math. Soc. (Ser. A)* **26** (1978), 257–269.

[2] R. FREUD and E. GYARMATI, *Number Theory*, Nemzeti Tankönyvkiadó, 2000 (in Hungarian).

[3] K. GYŐRY, Power values of products of consecutive integers and binomial coefficients, in *Number Theory and Its Applications* (S. Kanemitsu and K. Győry eds.), Kluwer Academic Publishers, 1999, 145–156.

[4] I. SCHUR, Einige Sätze über Primzahlen mit Anwendungen auf Irreduzibilitätsfragen, *S. B. Deutsch. Akad. Wiss. Berlin Kl. Math. Phys. Tech.* **23** (1929), 1–24.

[5] J. J. SYLVESTER, On arithmetical series, *Messenger Math.* **21** (1892), 1–19, 87–120. (*Collected Mathematical Papers* **vol. 4** (1912), 687–731.)

Gábor Nyul

University of Debrecen
Institute of Mathematics
H–4010 Debrecen Pf.12., Hungary
gnyul@math.klte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 123–126

CLOSURES OF OPEN SETS IN GENERALIZED TOPOLOGICAL SPACES

By

ÁKOS CSÁSZÁR*

(Received November 6, 2004)

1. Introduction and preliminaries

One of the generalizations of topologies is the *generalized topology* (briefly GT) in the sense of [3]. A GT on a set X is a subset μ of the power set $\exp X$ of X such that $\emptyset \in \mu$ and every union of elements of μ belongs to μ. If μ is a GT on X in the above sense, the elements of μ are said to be *μ-open* and their complements *μ-closed.* If $A \subset X$, we denote by $i_\mu A$ the union of all μ-open sets contained in A and by $c_\mu A$ the intersection of all μ-closed sets contining A. They clearly determine the largest μ-open set contained in A and the smallest μ-closed set containing A, respectively.

Let us call *set function* a map $\gamma\colon \exp X \to \exp X$. Then both i_μ and c_μ are set functions, *monotone* (i.e. $A \subset B \subset X$ implies $i_\mu A \subset i_\mu B$ and $c_\mu A \subset c_\mu B$) and *idempotent* (i.e. $i_\mu i_\mu A = i_\mu A$ and $c_\mu c_\mu A = c_\mu A$ for $A \subset X$) (see [4]).

Let us denote by Γ the collection of all monotone set functions on X. If $\gamma \in \Gamma$, we say that a set $A \subset X$ is γ-open iff $A \subset \gamma A$. Then the collection of all γ-open sets is a GT $\mu = \lambda_\gamma$, on X (see [2]). We often say the λ_γ-closed sets to be *γ-closed.*

According to [2], if $\gamma \in \Gamma$, we define the *conjugate* γ^* of γ by

$$\gamma^* A = X - \gamma(X - A). \tag{1}$$

Obviously $\gamma^* \in \Gamma$, in particular, $i_\mu^* = c_\mu$ for a GT μ (see [4]). Clearly $(\gamma^*)^* A = \gamma A$. A set $A \subset X$ is γ-closed iff $A \supset \gamma^* A$ (see [2]).

---

* Research supported by Hungarian Foundation for Scientific Research, grant No. T 032042.

If $\gamma, \gamma' \in \Gamma$, clearly $\gamma \circ \gamma' \in \Gamma$; we write for the sake of simplicity $\gamma\gamma'$ instead of $\gamma \circ \gamma'$. Then $(\gamma\gamma')^* = \gamma^*(\gamma')^*$ (see [2]).

Let us fix a GT μ on X. Then the collection of all $i_\mu c_\mu i_\mu$-open sets is denoted by α, that of all $c_\mu i_\mu$-open sets by σ, that of all $i_\mu c_\mu$-open sets by π, that of all $c_\mu i_\mu c_\mu$-open sets by β. In the literature, when μ is a topology, the elements of α are said to be *α-open* [7], those of σ are called *semi-open* [5], those of π *preopen* [6], those of β *β-open* [1].

The purpose of the present paper is to prove some theorems that assert the existence of many pairs (γ, γ') of monotone set functions such that the c_{λ_γ}-closure of a γ'-open set can be easily determined.

2. Closures of γ-open sets

The main theorem is easily obtained:

THEOREM 2.1. *If $\gamma \in \Gamma$ satisfies*

$$\gamma\gamma A \subset A \quad \text{for} \quad A \subset X \tag{2.1.1}$$

(in particular, if $\gamma \in \Gamma$ is idempotent), then

$$\gamma A = c_{\lambda_{\gamma^*}} A \quad \text{for} \quad A \subset X. \tag{2.1.2}$$

PROOF. (2.1.1) implies that γA is γ^*-closed since $(\gamma^*)^* = \gamma$. It contains A by hypothesis and is the smallest set with this property. In fact, if $F \supset A$ is γ^*-closed then $F \supset \gamma F \supset \gamma A$. Therefore $\gamma A = c_{\lambda_{\gamma^*}} A$. ∎

In order to obtain useful consequences of the above theorem, let us recall that, according to [4], a set function $\gamma_1 \ldots \gamma_n$ such that each γ_i equals either i_μ or c_μ where μ is an arbitrary GT, is always idempotent. So we can state, by fixing an arbitrary GT μ on X:

COROLLARY 2.2. *If $A \subset X$ belongs to α then*

$$i_\mu c_\mu i_\mu A = c_\beta A.$$

PROOF. Apply 2.1 for $\gamma = i_\mu c_\mu i_\mu$ that is idempotent. Observe that $\lambda_\gamma = \alpha$ and $\gamma^* = c_\mu i_\mu c_\mu$ so that $\lambda_{\gamma^*} = \beta$. ∎

COROLLARY 2.3. *If $A \in \sigma$ then*

$$c_\mu i_\mu A = c_\mu A = c_\pi A.$$

PROOF. We choose the idempotent set function $\gamma = c_\mu i_\mu$. Then $A \in \sigma$ implies

$$A \subset c_\mu i_\mu A \subset c_\mu A \subset c_\mu c_\mu i_\mu A = c_\mu i_\mu A$$

while $\gamma^* = i_\mu c_\mu$ so that $\lambda_{\gamma^*} = \pi$.

COROLLARY 2.4. *If $A \in \pi$ then*

$$i_\mu c_\mu A = c_\sigma A.$$

PROOF. We choose $\gamma = i_\mu c_\mu$ so that $\gamma^* = c_\mu i_\mu$.

COROLLARY 2.5. *If $A \in \beta$ then*

$$c_\mu i_\mu c_\mu A = c_\mu A = c_\alpha A.$$

PROOF. Let $\gamma = c_\mu i_\mu c_\mu$ so that $A \in \beta$ means that A is γ-open. Now

$$A \subset c_\mu i_\mu c_\mu A \subset c_\mu c_\mu A = c_\mu A \subset c_\mu c_\mu i_\mu c_\mu A = c_\mu i_\mu c_\mu A$$

shows that $\gamma A = c_\mu A$ whenever $A \in \beta$. Further $\gamma^* = i_\mu c_\mu i_\mu$ so that $\lambda_{\gamma^*} = \alpha$.

For the particular case when μ is a topology, 2.3 and 2.5 can be found in [8] and 2.4 in [9].

References

[1] M. E. ABD EL-MONSEF, S. N. EL-DEEB and R.A. MAHMOUD, β-open sets and β-continuous mappings, *Bull. Fac. Sci. Assiut Univ.,* **12** (1983), 77–90.

[2] Á. CSÁSZÁR, Generalized open sets, *Acta Math. Hungar.,* **75** (1997), 65–87.

[3] Á. CSÁSZÁR, Generalized topology, generalized continuity, *Acta Math. Hungar.,* **96** (2002), 351–357.

[4] Á. CSÁSZÁR, Generalized open sets in generalized topologies, *Acta Math. Hungar.*

[5] N. LEVINE, Semi-open sets and semi-continuity in topological spaces, *Amer. Math. Monthly,* **70** (1963), 36–41.

[6] A. S. MASHHOUR, I. A. HASANEIN and S. N. EL-DEEB, On precontinuous and weak precontinuous mappings, *Proc. Math. Phys. Soc. Egypt,* **53** (1982), 47–53.

[7] O. NJÅSTAD, On some classes of nearly open sets, *Pacific J. Math.,* **15** (1965), 961–970.

[8] T. NOIRI, On almost continuous functions, *Indian J. Pure Appl. Math.,* **20** (1989), 571–576.

[9] T. NOIRI, Almost quasi-continuous functions, *Bull. Inst. Math. Acad. Sinica,* **18** (1990), 321–332.

Ákos Császár

Department of Analysis,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
csaszar@ludens.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 127–137

ON CONTRA-CONTINUITY

By

ERDAL EKICI

(Received 23 November, 2004)

1. Introduction and preliminaries

It is clear that various types of continuity play significant role in several branches of mathematics. Continuity of functions is one of important and basic topics in the topology. The purpose of this paper is to define contra-γ-continuous functions and to obtain several characterizations and properties of contra-γ-continuous functions. Moreover, the relationships between contra-γ-continuous functions and several concepts are also discussed.

In this paper, spaces X and Y always mean topological spaces on which no separation axioms are assumed unless explicitly stated.

A subset A of a space X is said to be preopen [10] if $A \subset \operatorname{int}(\operatorname{cl}(A))$. The complement of a preopen set is said to be preclosed [5]. For a subset A of X, $\operatorname{cl}(A)$ and $\operatorname{int}(A)$ represent the closure of A with respect to τ and the interior of A with respect to τ, respectively.

A subset A of a space X is said to be regular open (respectively regular closed) if $A = \operatorname{int}(\operatorname{cl}(A))$ (respectively $A = \operatorname{cl}(\operatorname{int}(A))$) [14].

A subset A is said to be b-open [1] or γ-open [6] or sp-open [2] (resp. α-open [11], semi-open [9]), if $A \subset \operatorname{cl}(\operatorname{int}(A)) \cup \operatorname{int}(\operatorname{cl}(A))$ (resp. $A \subset$ $\subset \operatorname{int}(\operatorname{cl}(\operatorname{int}(A)))$, $A \subset \operatorname{cl}(\operatorname{int}(A))$). The complement of a semi-open set is said to be semi-closed.

The complement of a γ-open set is said to be γ-closed [6]. The intersection of all γ-closed sets of X containing A is called the γ-closure [6] of A

Mathematics Subject Classifications: 54C08, 54C10, 54C05

and is denoted by $\gamma - \mathrm{cl}(A)$. The union of all γ-open sets of X contained A is called γ-interior of A and is denoted by $\gamma - \mathrm{int}(A)$.

The family of all γ-open (resp. α-open, semi-open, γ-closed, closed) sets of X is denoted by $\gamma O(X)$ (resp. $\alpha O(X)$, $SO(X)$, $\gamma C(X)$, $C(X)$).

The family of all γ-open (resp. γ-closed, closed) sets of X containing a point $x \in X$ is denoted by $\gamma O(X, x)$ (resp. $\gamma C(X, x)$, $C(X, x)$).

2. Characterizations

In this section, several properties of contra-γ-continuous functions are studied.

DEFINITION 1. A function $f : X \to Y$ is called contra-γ-continuous at a point $x \in X$ if for each closed set V in Y with $f(x) \in V$, there exists a γ-open set U in X such that $x \in U$ and $f(U) \subset V$ and f is called contra-γ-continuous if it has this property at each point of X.

THEOREM 2. *The following are equivalent for a function $f : X \to Y$:*

(1) f is contra-γ-continuous;

(2) the inverse image of a closed set of Y is γ-open;

(3) the inverse image of an open set of Y is γ-closed.

PROOF. (1)⇒(2): Let V be a closed set in Y with $x \in f^{-}(V)$. Since $f(x) \in V$ and f is contra-γ-continuous, there exists a γ-open set U in X containing x such that $f(U) \subset V$. It follows that $x \in U \subset f^{-1}(V)$. Hence $f^{-1}(V)$ is γ-open.

(2)⇒(3): Let U be any open set of Y. Since $Y \setminus U$ is closed, then by (2), it follows that $f^{-1}(Y \setminus U) = X \setminus f^{-1}(U)$ is γ-open. This shows that $f^{-1}(U)$ is γ-closed in X.

(3)⇒(1): Let $x \in X$ and V be a closed set in Y with $f(x) \in V$. By (3), it follows that $f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$ is γ-closed and so $f^{-1}(V)$ is γ-open. Take $U = f^{-1}(V)$. We obtain that $x \in U$ and $f(U) \subset V$. This shows that f is contra-γ-continuous. ∎

DEFINITION 3. A function $f : X \to Y$ is said to be

(1) contra-continuous [3] if $f^{-1}(V)$ is closed in X for every open set V of Y,

(2) RC-continuous [4] if $f^{-1}(V)$ is regular closed in X for each open set V of Y,

(3) contra-precontinuous [7] if $f^{-1}(V)$ is preclosed in X for each open set V of Y,

(4) contra-semicontinuous [4] if $f^{-1}(V)$ is semi-closed in X for each open set V of Y.

REMARK 4. The following diagram holds:

$$\begin{array}{ccc} RC\text{-continuous} & & \\ \Downarrow & & \\ \text{contra-continuous} & \Rightarrow & \text{contra-semicontinuous} \\ \Downarrow & & \Downarrow \\ \text{contra-precontinuous} & \Rightarrow & \text{contra-}\gamma\text{-continuous} \end{array}$$

None of these implications is reversible.

EXAMPLE 5. Consider the set $\mathbb{R}$ of real numbers with the usual topology τ_u and let $S = [0,1] \cup ((1,2) \cap \mathbb{Q})$ where $\mathbb{Q}$ stands for the set of rational numbers. Then S is γ-open but neither semi-open nor preopen.

Let $f : (\mathbb{R}, \tau_u) \to (\mathbb{R}, \tau_D)$ be a identity function where τ_D is discrete topology on $\mathbb{R}$.

S is closed in $(\mathbb{R}, \tau_D)$. $f^{-1}(S) = S$ is γ-open in $(\mathbb{R}, \tau_u)$ but not preopen and not semi-open. Hence, it is obtained that f is contra-γ-continuous but not contra-precontinuous function and not contra-semicontinuous function.

The other implications are not reversible as shown in [4, 7].

THEOREM 6. *Let $f : X \to Y$ be a function and let $g : X \to X \times Y$ be the graph function of f, defined by $g(x) = (x, f(x))$ for every $x \in X$. If g is contra-γ-continuous, then f is contra-γ-continuous.*

PROOF. Let U be an open set in Y, then $X \times U$ is an open set in $X \times Y$. It follows from Theorem 2 that $f^{-1}(U) = g^{-1}(X \times U) \in \gamma C(X)$. Thus, f is contra-γ-continuous. ∎

DEFINITION 7. A filter base Λ is said to be γ-convergent (resp. c-convergent) to a point x in X if for any $U \in \gamma O(X)$ containing x (resp. $U \in C(X)$ containing x), there exists a $B \in \Lambda$ such that $B \subset U$.

THEOREM 8. *If a function $f : X \to Y$ is contra-γ-continuous, then for each point $x \in X$ and each filter base Λ in X which is γ-convergent to x, the filter base $f(\Lambda)$ is c-convergent to $f(x)$.*

PROOF. Let $x \in X$ and Λ be any filter base in X which is γ-convergent to x. Since f is contra-γ-continuous, then for any $V \in C(Y)$ containing $f(x)$, there exists $U \in \gamma O(X)$ containing x such that $f(U) \subset V$. Since Λ is γ-convergent to x, there exists a $B \in \Lambda$ such that $B \subset U$. This means that $f(B) \subset V$ and therefore the filter base $f(\Lambda)$ is c-convergent to $f(x)$. ∎

LEMMA 9. *Let A and X_0 be subsets of a space (X,τ). If $A \in \gamma O(X)$ and $X_0 \in \alpha O(X)$, then $A \cap X_0 \in \gamma O(X_0)$ [1, 6].*

LEMMA 10. *Let $A \subset X_0 \subset X$, $A \in \gamma O(X_0)$ and $X_0 \in \alpha O(X)$, then $A \in \gamma O(X)$ [6].*

LEMMA 11. *The intersection of an open and a γ-open set is a γ-open set [4].*

DEFINITION 12. A function $f : X \to Y$ is called weakly continuous if for each $x \in X$ and each open set G containing $f(x)$, there exists an open set U in X containing x such that $f(U) \subset \mathrm{cl}(G)$ [8].

THEOREM 13. *If $f : X \to Y$ is weakly continuous, $g : X \to Y$ is contra-γ-continuous and Y is Urysohn, then $E = \{x \in X : f(x) = g(x)\}$ is γ-closed in X.*

PROOF. If $x \in X \setminus E$, then it follows that $f(x) \neq g(x)$. Since Y is Urysohn, there exist open sets V and W such that $f(x) \in V$, $g(x) \in W$ and $\mathrm{cl}(V) \cap \cap\, \mathrm{cl}(W) = \emptyset$. Since f is weakly continuous and g is contra-γ-continuous, there exist an open set U containing x and a γ-open set G containing x such that $f(U) \subset \mathrm{cl}(V)$ and $g(G) \subset \mathrm{cl}(W)$. Set $O = U \cap G$. By the previous lemma, O is γ-open in X. Therefore $f(O) \cap g(O) = \emptyset$ and it follows that $x \notin \gamma - \mathrm{cl}(E)$. This shows that E is γ-closed in X. ∎

THEOREM 14. *Let $f : X \to Y$ be a function and $x \in X$. If there exists $U \in \alpha O(X)$ such that $x \in U$ and the restriction of f to U is a contra-γ-continuous function at x, then f is contra-γ-continuous at x.*

PROOF. Suppose that $F \in C(Y)$ containing $f(x)$. Since $f\,|_U$ is contra-γ-continuous at x, there exists $V \in \gamma O(U)$ containing x such that $f(V) = = (f\,|_U)(V) \subset F$. Since $U \in \alpha O(X)$ containing x, it follows from Lemma 10 that $V \in \gamma O(X)$ containing x. This shows clearly that f is contra-γ-continuous at x. ∎

THEOREM 15. *If $f : X \to Y$ is a contra-γ-continuous function and A is any α-open subset of X, then the restriction $f|_A : A \to Y$ is contra-γ-continuous.*

PROOF. Let F be a closed set in Y. Then, by Theorem 2, $f^{-1}(F) \in \in \gamma O(X)$. Since A is α-open in X, it follows from Lemma 9 that $(f|_A)^{-1}(F) = A \cap f^{-1}(F) \in \gamma O(A)$. Therefore, $f|_A$ is a contra-γ-continuous function. ∎

THEOREM 16. *Let $f : X \to Y$ be a function and $\{U_i : i \in I\}$ be a α-open cover of X. If for each $i \in I$, $f|_{U_i}$ is contra-γ-continuous, then $f : X \to Y$ is a contra-γ-continuous function.*

PROOF. Let F be a closed set in Y. Since $f|_{U_i}$ is contra-γ-continuous for each $i \in I$, $(f|_{U_i})^{-1}(F) \in \gamma O(U_i)$. Since $U_i \in \alpha O(X)$, by the Lemma 10, $(f|_{U_i})^{-1}(F) \in \gamma O(X)$ for each $i \in I$. Then $f^{-1}(F) = \bigcup_{i \in I}[(f|_{U_i})^{-1}(F)] \in \in \gamma O(X)$. This gives f is a contra-γ-continuous function. ∎

DEFINITION 17. A function $f : X \to Y$ is said to be γ-irresolute if for each $x \in X$ and each $V \in \gamma O(Y, f(x))$, there exists a γ-open set U in X containing x such that $f(U) \subset V$.

THEOREM 18. *Let $f : X \to Y$ and $g : Y \to Z$ be functions. Then, the following properties hold:*

(1) If f is γ-irresolute and g is contra-γ-continuous, then $g \circ f : X \to Z$ is contra-γ-continuous.

(2) If f is contra-γ-continuous and g is continuous, then $g \circ f : X \to Z$ is contra-γ-continuous.

PROOF. (1) Let $x \in X$ and $W \in C(Z, (g \circ f)(x))$. Since g is contra-γ-continuous, there exists a γ-open set V in Y containing $f(x)$ such that $g(V) \subset W$. Since f is γ-irresolute, there exists a γ-open set U in X containing x such that $f(U) \subset V$. This shows that $(g \circ f)(U) \subset W$. Therefore, $g \circ f$ is contra-γ-continuous.

(2) Let $x \in X$ and $W \in C(Z, (g \circ f)(x))$. Since g is continuous, $V = g^{-1}(W)$ is closed. Since f is contra-γ-continuous, there exists a γ-open set U in X containing x such that $f(U) \subset V$. Hence $(g \circ f)(U) \subset W$. This shows that $g \circ f$ is contra-γ-continuous. ∎

DEFINITION 19. A function $f : X \to Y$ is called γ-open if image of each γ-open set is γ-open.

THEOREM 20. *If $f : X \to Y$ is a surjective γ-open function and $g : Y \to Z$ is a function such that $g \circ f : X \to Z$ is contra-γ-continuous, then g is contra-γ-continuous.*

PROOF. Suppose that x and y are two points in X and Y, respectively, such that $f(x) = y$. Let $V \in C(Z, (g \circ f)(x))$. Then there exists a γ-open set U in X containing x such that $g(f(U)) \subset V$. Since f is γ-open, $f(U)$ is a γ-open set in Y containing y such that $g(f(U)) \subset V$. This implies that g is contra-γ-continuous. ∎

COROLLARY 21. *Let $f : X \to Y$ be a surjective γ-irresolute and γ-open function and let $g : Y \to Z$ be a function. Then, $g \circ f : X \to Z$ is contra-γ-continuous if and only if g is contra-γ-continuous.*

PROOF. It can be obtained from Theorem 18 and Theorem 20. ∎

DEFINITION 22. A function $f : X \to Y$ is called weakly contra-γ-continuous if for each $x \in X$ and each closed set F of Y containing $f(x)$, there exists a γ-open set U in X containing x such that $\mathrm{int}(f(U)) \subset V$.

DEFINITION 23. A function $f : X \to Y$ is called γ-semi-open if image of each γ-open set is semi-open.

THEOREM 24. *If a function $f : X \to Y$ is weakly contra-γ-continuous and γ-semi-open, then f is contra-γ-continuous.*

PROOF. Let $x \in X$ and F be a closed set containing $f(x)$. Since f is weakly contra-γ-continuous, there exists a γ-open set U in X containing x such that $\mathrm{int}(f(U)) \subset F$. Since f is γ-semi-open, $f(U) \in SO(Y)$ and $f(U) \subset \mathrm{cl}(\mathrm{int}(f(U))) \subset F$. This shows that f is contra-γ-continuous. ∎

3. Some properties

In this section, graphs and preservation theorems of contra-γ-continuity are investigated.

DEFINITION 25. A space X is called γ-connected provided that X is not the union of two disjoint nonempty γ-open sets [6].

THEOREM 26. *If $f : X \to Y$ is contra-γ-continuous surjection and X is γ-connected, then Y is connected.*

PROOF. Suppose that Y is not connected space. There exist nonempty disjoint open sets V_1 and V_2 such that $Y = V_1 \cup V_2$. Therefore, V_1 and V_2 are clopen in Y. Since f is contra-γ-continuous, $f^{-1}(V_1)$ and $f^{-1}(V_2)$ are γ-open in X. Moreover, $f^{-1}(V_1)$ and $f^{-1}(V_2)$ are nonempty disjoint and $X = f^{-1}(V_1) \cup f^{-1}(V_2)$. This shows that X is not γ-connected. This contradicts that Y is not connected assumed. Hence, Y is connected. ∎

DEFINITION 27. A topological space is called γ-ultra-connected if every two non-void γ-closed subsets of X intersect and is called hyperconnected [13] if every open set is dense.

THEOREM 28. *If X is γ-ultra-connected and $f : X \to Y$ is contra-γ-continuous and surjective, then Y is hyperconnected.*

PROOF. Assume that Y is not hyperconnected. Then there exists an open set V such that V is not dense in Y. Then there exist disjoint non-empty open subsets B_1 and B_2 in Y, namely $\mathrm{int}(\mathrm{cl}(V))$ and $Y \setminus \mathrm{cl}(V)$. Since f is contra-γ-continuous and onto, by Theorem 2, $A_1 = f^{-1}(B_1)$ and $A_2 = = f^{-1}(B_2)$ are disjoint non-empty γ-closed subsets of X. By assumption, the γ-ultra-connectedness of X implies that A_1 and A_2 must intersect. By contradiction, Y is hyperconnected. ∎

DEFINITION 29. A space X is said to be

(1) weakly Hausdorff [12] if each element of X is an intersection of regular closed sets,

(2) γ-Hausdorff if for each pair of distinct points x and y in X, there exist $U \in \gamma O(X, x)$ and $V \in \gamma O(X, y)$ such that $U \cap V = \emptyset$,

(3) γ-T_1 if for each pair of distinct points in X, there exist γ-open sets U and V containing x and y, respectively, such that $y \notin U$ and $x \notin V$.

Recall that for a function $f : X \to Y$, the subset $\{(x, f(x)) : x \in X\} \subset$ $\subset X \times Y$ is called the graph of f and is denoted by $G(f)$.

DEFINITION 30. A graph $G(f)$ of a function $f : X \to Y$ is said to be contra-γ-closed if for each $(x, y) \in (X \times Y) \backslash G(f)$, there exist a γ-open set U in X containing x and $V \in C(Y, y)$ such that $(U \times V) \cap G(f) = \emptyset$.

LEMMA 31. *The following properties are equivalent for a graph $G(f)$ of a function f:*

(1) $G(f)$ is contra-γ-closed;

(2) for each $(x, y) \in (X \times Y) \backslash G(f)$, there exist a γ-open set U in X containing x and $V \in C(Y, y)$ such that $f(U) \cap V = \emptyset$.

PROOF. Obvious. ∎

THEOREM 32. *If $f : X \to Y$ is contra-γ-continuous and Y is Urysohn, $G(f)$ is contra-γ-closed graph in $X \times Y$.*

PROOF. Suppose that Y is Urysohn. Let $(x, y) \in (X \times Y) \backslash G(f)$. It follows that $f(x) \neq y$. Since Y is Urysohn, there exist open sets V and W such that $f(x) \in V$, $y \in W$ and $\mathrm{cl}(V) \cap \mathrm{cl}(W) = \emptyset$. Since f is contra-γ-continuous, there exists a γ-open set U in X containing x such that $f(U) \subset \mathrm{cl}(V)$. Therefore, $f(U) \cap \mathrm{cl}(W) = \emptyset$ and $G(f)$ is contra-γ-closed in $X \times Y$. ∎

THEOREM 33. *Let $f : X \to Y$ have a contra-γ-closed graph. If f is injective, then X is γ-T_1.*

PROOF. Let x and y be any two distinct points of X. Then, we have $(x, f(y)) \in (X \times Y) \backslash G(f)$. By Lemma 31, there exist a γ-open set U in X containing x and $F \in C(Y, f(y))$ such that $f(U) \cap F = \emptyset$; hence $U \cap f^{-1}(F) = \emptyset$. Therefore, we have $y \notin U$. This implies that X is γ-T_1. ∎

DEFINITION 34. A space X said to be

(1) γ-compact [6] (strongly S-closed [3]) if every γ-open (respectively closed) cover of X has a finite subcover.

(2) countably γ-compact (strongly countably S-closed) if every countable cover of X by γ-open (respectively closed) sets has a finite subcover.

(3) γ-Lindelof (strongly S-Lindelof) if every γ-open (respectively closed) cover of X has a countable subcover.

THEOREM 35. *Contra-γ-continuous images of γ-compact (γ-Lindelof, countably γ-compact) spaces are strongly S-closed (respectively strongly S-Lindelof, strongly countably S-closed).*

PROOF. Suppose that $f : X \to Y$ is a contra-γ-continuous surjection. Let $\{V_\alpha : \alpha \in I\}$ be any closed cover of Y. Since f is contra-γ-continuous, then $\{f^{-1}(V_\alpha) : \alpha \in I\}$ is a γ-open cover of X and hence there exists a finite subset I_0 of I such that $X = \bigcup\{f^{-1}(V_\alpha) : \alpha \in I_0\}$. Therefore, we have $Y = \bigcup\{V_\alpha : \alpha \in I_0\}$ and Y is strongly S-closed.

The other proofs can be obtained similarly. ∎

DEFINITION 36. A space X said to be (1) γ-closed-compact if every γ-closed cover of X has a finite subcover, (2) countably γ-closed-compact if every countable cover of X by γ-closed sets has a finite subcover, (3) γ-closed-Lindelof if every cover of X by γ-closed sets has a countable subcover.

THEOREM 37. *contra-γ-continuous images of γ-closed-compact (γ-closed-Lindelof, countably γ-closed-compact) spaces are compact (respectively Lindelof, countably compact).*

PROOF. Suppose that $f : X \to Y$ is a contra-γ-continuous surjection. Let $\{V_\alpha : \alpha \in I\}$ be any open cover of Y. Since f is contra-γ-continuous, then $\{f^{-1}(V_\alpha) : \alpha \in I\}$ is a γ-closed cover of X. Since X is γ-closed-compact, there exists a finite subset I_0 of I such that $X = \bigcup\{f^{-1}(V_\alpha) : \alpha \in I_0\}$. Thus, we have $Y = \bigcup\{V_\alpha : \alpha \in I_0\}$ and Y is compact.

The other proofs can be obtained similarly. ∎

THEOREM 38. *If f is a contra-γ-continuous injection and Y is Urysohn, then X is γ-Hausdorff.*

PROOF. Suppose that Y is Urysohn. By the injectivity of f, it follows that $f(x) \neq f(y)$ for any distinct points x and y in X. Since Y is Urysohn, there exist open sets V and W such that $f(x) \in V$, $f(y) \in W$ and $\mathrm{cl}(V) \cap \cap\, \mathrm{cl}(W) = \emptyset$. Since f is a contra-γ-continuous, there exist γ-open sets U and G in X containing x and y, respectively, such that $f(U) \subset \mathrm{cl}(V)$ and $f(G) \subset \mathrm{cl}(W)$. Hence $U \cap G = \emptyset$. This shows that X is γ-Hausdorff. ∎

THEOREM 39. *If f is a contra-γ-continuous injection and Y is weakly Hausdorff, then X is γ-T_1.*

PROOF. Suppose that Y is weakly Hausdorff. For any distinct points x and y in X, there exist regular closed sets V, W in Y such that $f(x) \in V$, $f(y) \notin V$, $f(x) \notin W$ and $f(y) \in W$. Since f is contra-γ-continuous, by Theorem 2, $f^{-1}(V)$ and $f^{-1}(W)$ are γ-open subsets of X such that $x \in f^{-1}(V)$, $y \notin f^{-1}(V)$, $x \notin f^{-1}(W)$ and $y \in f^{-1}(W)$. This shows that X is γ-T_1. ∎

References

[1] D. ANDRIJEVIC, On b-open sets, *Mat. Vesnik* **48** (1996), 59–64.

[2] J. DONTCHEV and M. PRZEMSKI, On the various decompositions of continuous and some weakly continuous functions, *Acta Math. Hungar.* **71(1–2)** (1996), 109–120.

[3] J. DONTCHEV, Contra-continuous functions and strongly S-closed spaces, *Internat. J. Math. Math. Sci.* **19(2)** (1996), 303–310.

[4] J. DONTCHEV and T. NOIRI, Contra-semicontinuous functions, *Math. Pannonica* **10** (1999), 159–168.

[5] S. N. EL-DEEB, I. A. HASANEIN, A. S. MASHHOUR and T. NOIRI, On p-regular spaces, *Bull. Math. Soc. Sci. Math. R. Soc. Roumania* **27(75)** (1983), 311–315.

[6] A. A. EL-ATIK, *A study of some types of mappings on topological spaces*, Master's Thesis, Faculty of Science, Tanta University, Tanta, Egypt, 1997.

[7] S. JAFARI and T. NOIRI, On contra-precontinuous functions, *Bulletin of the Malaysian Mathematical Society* **25(1)** (2002).

[8] N. LEVINE, A decomposition of continuity in topological spaces, *Amer. Math. Monhly* **68** (1961), 44–46.

[9] N. LEVINE, Semi-open sets and semi-continuity in topological spaces, *Amer. Math. Monthly* **70** (1963), 36–41.

[10] A. S. MASHHOUR, M. E. ABD EL-MONSEF and S. N. EL-DEEB, On precontinuous and weak precontinuous mappings, *Proc. Phys. Soc. Egypt* **53** (1982), 47–53.

[11] O. NJASTAD, On some classes of nearly open sets, *Pasific J. Math.* **15** (1965), 961–970.

[12] T. SOUNDARARAJAN, Weakly Hausdorff spaces and the cardinality of topological spaces, in: *General Topology and its Relation to Modern Analysis and Algebra*. III, Proc. Conf. Kanpur, 1968, Academia, Prague 1971, pp. 301–306.

[13] L. A. STEEN and J. A. SEEBACH JR, *Counterexamples in Topology*, Holt, Rinerhart and Winston, New York, 1970.

[14] M. H. STONE, Applications of the theory of Boolean rings to general topology, *TAMS* **41** (1937), 375–381.

Erdal Ekici

Department of Mathematics
Canakkale Onsekiz Mart University
Terzioglu Campus
17020 Canakkale
TURKEY
eekici@comu.edu.tr

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 139–156

ON SYSTEMS OF NONLINEAR PARABOLIC FUNCTIONAL DIFFERENTIAL EQUATIONS

By

ÁDÁM BESENYEI

(Received December 2, 2004)

1. Introduction

Second order quasilinear parabolic differential equations where also the main part contains functional dependence on the unknown functions were studied e.g. in [1] by L. Simon. There the following equation was considered:

$$D_t u(t,x) - \sum_{i=1}^{n} D_i [a_i(t,x,u(t,x),Du(t,x);u)] + a_0(t,x,u(t,x),Du(t,x);u) =$$

$$= f(t,x) \qquad (t,x) \in Q_T = (0,T)\times\Omega, \quad a_i\colon Q_T \times \mathbb{R}^{n+1} \times L^p(0,T;V) \to \mathbb{R},$$

where V denotes a closed linear subset of the Sobolev-space $W^{1,p}(\Omega)$ $(2 \leq \leq p < \infty)$.

Let us now consider a system of this type of equations:

(1)

$$D_t u^{(l)}(t,x) - \sum_{i=1}^{n} D_i \left[a_i^{(l)} \left(t,x,u^{(1)}(t,x),\dots,u^{(N)}(t,x),\right.\right.$$

$$\left.\left. Du^{(1)}(t,x),\dots,Du^{(N)}(t,x);u^{(1)},\dots,u^{(N)}\right)\right] +$$

$$+ a_0^{(l)} \left(t,x,u^{(1)}(t,x),\dots,u^{(N)}(t,x),\right.$$

$$\left. Du^{(1)}(t,x),\dots,Du^{(N)}(t,x);u^{(1)},\dots,u^{(N)}\right) =$$

$$= F^{(l)}(t,x), \qquad (t,x) \in Q_T = (0,T)\times\Omega, \quad \Omega \subset \mathbb{R}^n, \quad l = 1,\dots,N.$$

In the next section we define the weak form of the above system and formulate conditions on the coefficients. With these we can prove existence of weak solutions. The conditions are generalizations of the classical Léray–Lions conditions for systems with some special conditions for these type of systems. Finally we show some examples.

2. Existence of weak solutions

First we introduce some notations. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with the C^1 regularity property and $2 \leq p < \infty$ be a real number. Denote by $W^{1,p}(\Omega)$ the usual Sobolev space of real valued functions with the norm

$$\|u\| = \left(\int_\Omega (|Du|^p + |u|^p) \right)^{\frac{1}{p}} .$$

Let $V_l \subset W^{1,p}(\Omega)$ $(l = 1, \dots, N)$ be a closed linear subspace (e.g. $W_0^{1,p}(\Omega)$ or $W^{1,p}(\Omega)$) and let $V = V_1 \times \dots \times V_N$. Denote by $L^p(0,T;V)$ the Banach space of measurable functions $u\colon (0,T) \to V$ such that $\|u\|^p$ is integrable and define the norm by

$$\|u\|_{L^p(0,T;V)} = \int_0^T \|u(t)\|_V^p \, dt.$$

The dual space of $L^p(0,T;V)$ is $L^q(0,T;V^*)$ where $\frac{1}{p} + \frac{1}{q} = 1$ and V^* is the dual space of V. Let $X = L^p(0,T;V)$ and $Y = L^p(0,T;(L^p(\Omega))^N)$. For $u \in$ $\in X$ we shall write $u = (u^{(1)}, \dots, u^{(N)})$, where $u^{(l)} \in L^p(0,T;V_l)$. A vector $\xi \in \mathbb{R}^{(n+1)N}$ is written in the form $\xi = (\zeta_0, \zeta)$, where $\zeta_0 = (\zeta_0^{(1)}, \dots, \zeta_0^{(N)}) \in$ $\in \mathbb{R}^N$ and $\zeta = (\zeta^{(1)}, \dots, \zeta^{(N)}) \in \mathbb{R}^{nN}$. Here $\zeta^{(l)} = (\zeta_1^{(l)}, \dots, \zeta_n^{(l)}) \in \mathbb{R}^n$. Now we formulate 5 essential assumptions on functions $a_i^{(l)}$ $(i = 0, \dots, n;$ $l = 1, \dots, N)$, which (as we will see) are sufficient for existence of weak solutions.

F1. Suppose that $a_i^{(l)}\colon Q_T \times \mathbb{R}^{(n+1)N} \times L^p(0,T;V) \to \mathbb{R}$ are Carathéodory functions for each $v \in L^p(0,T;V)$. This means that they are measurable in (t,x) for every $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$, and continuous in (ζ_0,ζ) for almost every $(t,x) \in Q_T$ $(i = 0, \dots, n;\ l = 1, \dots, N)$.

F2. Suppose that there exist bounded operators $g_1 : L^p(0,T;V) \to \mathbb{R}^+$ and $k_1 : L^p(0,T;V) \to L^q(Q_T)$ such that

$$|a_i^{(l)}(t,x,\zeta_0,\zeta;v)| \le g_1(v)\left(|\zeta_0|^{p-1} + |\zeta|^{p-1}\right) + \left[k_1(v)\right](t,x)$$

for a.e. $(t,x) \in Q_T$, each $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$ and $v \in L^p(0,T;V)$ $(i = 0,\dots,n;\ l = 1,\dots,N)$.

F3. Suppose that for each $\zeta \neq \eta \in \mathbb{R}^{nN}$, a.e. $(t,x) \in Q_T$, each $\zeta_0 \in \mathbb{R}^N$ and each $v \in L^p(0,T;V)$

$$\sum_{l=1}^{N}\sum_{i=1}^{n}\left(a_i^{(l)}(t,x,\zeta_0,\zeta;v) - a_i^{(l)}(t,x,\zeta_0,\eta;v)\right)(\zeta_i^{(l)} - \eta_i^{(l)}) > 0.$$

F4. Suppose that there exist operators $g_2 : L^p(0,T;V) \to \mathbb{R}^+$ and $k_2 : L^p(0,T;V) \to L^1(Q_T)$ such that

$$\sum_{l=1}^{N}\sum_{i=0}^{n} a_i^{(l)}(t,x,\zeta_0,\zeta;v)\zeta_i^{(l)} \ge g_2(v)\left(|\zeta_0|^p + |\zeta|^p\right) - [k_2(v)](t,x)$$

for a.e. $(t,x) \in Q_T$, each $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$ and $v \in L^p(0,T;V)$ $(i = 0,\dots,n;\ l = 1,\dots,N)$. Further, operators g_2, k_2 has the following property:

$$\lim_{\|v\|_{L^p(0,T;V)}\to\infty}\left(g_2(v)\|v\|^{p-1}_{L^p(0,T;V)} - \frac{\|k_2(v)\|_{L^1(Q_T)}}{\|v\|_{L^p(0,T;V)}}\right) = +\infty.$$

F5. Suppose that if $u_k \to u$ weakly in $L^p(0,T;V)$ and strongly in $L^p(0,T;(L^p(\Omega))^N)$, then

$$\lim_{k\to\infty}\|a_i^{(l)}(\cdot\,,u_k(\cdot),Du_k(\cdot);u_k) - a_i^{(l)}(\cdot\,,u_k(\cdot),Du_k(\cdot);u)\|_{L^q(Q_T)} = 0.$$

$(i = 0,\dots,n;\ l = 1,\dots,N)$

We now define the weak form of system (1). Let us introduce first the operator $A : L^p(0,T;V) \to L^q(0,T;V^*)$. For $u = (u^{(1)},\dots,u^{(N)}) \in L^p(0,T;V)$ and $v = (v^{(1)},\dots,v^{(N)}) \in L^p(0,T;V)$ define

$$[A(u),v] :=$$

$$\sum_{l=1}^{N}\int_{Q_T}\left[\sum_{i=1}^{n} a_i^{(l)}\left(t,x,u(t,x),Du(t,x);u\right)D_i v^{(l)}(t,x) +\right.$$

$$+ a_0^{(l)}(t,x,u(t,x),Du(t,x);u)\Big) v^{(l)}(t,x)\Big] \, dt dx,$$

where D_i denotes the operator of (distributional) partial differentiating with respect to x_i and $D = (D_1, \ldots, D_N)$. As usual let $L: L^p(0, T; V) \to L^q(0, T; V^*)$ be the following operator:

$$D(L) = \{u \in X : D_t u \in X^*, u(0) = 0\}, \qquad Lu = D_t u.$$

With operator A we define the weak form of system (1) by

$$D_t u + A(u) = F.$$

In the next theorem we prove some important properties of A from which existence of weak solution follows.

THEOREM 1. *Assume that conditions F1–F5 are fulfilled. Then $A: X \to X^*$ is bounded, demicontinuous, coercive and pseudomonotone with respect to $D(L)$.*

PROOF. The proof is based on elementary techincs and on Hölder's inequality.

BOUNDEDNESS. From triangle inequality it is clear that it is sufficient to deal with only one integral in $[A(u), v]$. This can be estimated by Hölder's inequality:

$$\left| \int_{Q_T} a_i^{(l)}(t,x,u(t,x),Du(t,x);u) D_i v^{(l)}(t,x) dt dx \right| \leq \tag{2}$$

$$\leq \left(\int_{Q_T} |a_i^{(l)}(t,x,u(t,x),Du(t,x);u)|^q dt dx \right)^{\frac{1}{q}} \left(\int_{Q_T} |D_i v^{(l)}(t,x)|^p dt dx \right)^{\frac{1}{p}}.$$

(In case $i = 0$ we replace $D_i v^{(l)}$ by $v^{(l)}$.) On the right hand side of (2) the second term is less or equal than $\|v\|_X$ and the first term can be estimated by the inequality $|a+b|^r \leq 2^{r-1} \cdot (|a|^r + |b|^r)$:

$$\left(\int_{Q_T} |a_i^{(l)}(t,x,u(t,x),Du(t,x);u)|^q dt dx \right)^{\frac{1}{q}} \leq \tag{3}$$

$$\leq const \cdot \left(\int_{Q_T} \left[g_1(u)^q \left(|u(t,x)|^{(p-1)q} + |Du(t,x)|^{(p-1)q} \right) + \right. \right.$$

$$+ |[k_1(u)](t,x)|^q \Big] \, dt dx \Bigg)^{\frac{1}{q}} \le$$

$$\le const \cdot \left[g_1(u) \left(\int_{Q_T} |u|^p + |Du|^p \right)^{\frac{1}{q}} + \left(\int_{Q_T} |k_1(u)|^q \right)^{\frac{1}{q}} \right] =$$

$$= const \cdot \left(g_1(u) \|u\|_X^{\frac{p}{q}} + \|k_1(u)\|_{L^q(Q_T)} \right).$$

Summing the above estimations with respect to i and l we get:

$$|[A(u), v]| \le const \cdot \left(g_1(u) \|u\|_X^{\frac{p}{q}} + \|k_1(u)\|_{L^q(Q_T)} \right) \|v\|_X.$$

This means that $\|A(u)\|_{X^*} \le const \cdot \left(g_1(u) \|u\|_X^{\frac{p}{q}} + \|k_1(u)\|_{L^q(Q_T)} \right)$. From here by boundedness of operators g_1 and k_1 follows the boundedness of A.

DEMICONTINUITY. Assume that $u_k \to u$ strongly in X. Then there exists a subsequence $(\tilde{u}_k) \subset (u_k)$, such that $\tilde{u}_k \to u$ and $D\tilde{u}_k \to Du$ for a.e. $(t,x) \in Q_T$. We show that for each $v \in X$ we have $[A(\tilde{u}_k) - A(u), v] \to 0$, then using the subsequence trick the proof of demicontinuity will be finished. It is useful to introduce operator $\tilde{A}_u \colon X \to X^*$ (u is fixed) defined by

$$[\tilde{A}_u(v), w] := \sum_{l=1}^{N} \int_{Q_T} \left[\sum_{i=1}^{n} a_i^{(l)}(t,x,v(t,x), Dv(t,x); u) D_i w^{(l)}(t,x) + \right.$$

$$\left. + a_0^{(l)}(t,x,v(t,x), Dv(t,x); u) w^{(l)}(t,x) \right] dt dx.$$

We prove that $A(\tilde{u}_k) - \tilde{A}_u(\tilde{u}_k) \to 0$ and $\tilde{A}_u(\tilde{u}_k) - A(u) \to 0$ weakly in X^*. It is easy to see (from triangle and Hölder's inequality) that it is sufficient to show

$$\|a_i^{(l)}(\cdot\,, \tilde{u}_k(\cdot), D\tilde{u}_k(\cdot); \tilde{u}_k) - a_i^{(l)}(\cdot\,, \tilde{u}_k(\cdot), D\tilde{u}_k(\cdot); u)\|_{L^q(Q_T)} \to 0 \tag{4}$$

and

$$\|a_i^{(l)}(\cdot\,, \tilde{u}_k(\cdot), D\tilde{u}_k(\cdot); u) - a_i^{(l)}(\cdot\,, u(\cdot), Du(\cdot); u)\|_{L^q(Q_T)} \to 0. \tag{5}$$

The strong convergence in X implies the weak convergence in X, and because of the continuous imbedding $X \to Y$ it implies the weak convergence in Y, too. So that from F5 it follows that (4) is true indeed. On the other hand, from condition F1 we know that $a_i^{(l)}$ is continuous in (ζ_0, ζ), hence

$$a_i^{(l)}(t,x,\tilde{u}_k(t,x), D\tilde{u}_k(t,x);u) \to a_i^{(l)}(t,x,u(t,x), Du(t,x);u)$$

for a.e. $(t,x) \in Q_T$, by the almost everywhere convergence of $\tilde{u}_k$ and $D\tilde{u}_k$ in Q_T. Further,

$$|a_i^{(l)}(t,x,\tilde{u}_k(t,x), D\tilde{u}_k(t,x);u)|^q \le$$
$$\le g_1(u)^q \left(|\tilde{u}_k(t,x)|^p + |D\tilde{u}_k(t,x)|^p\right) + |[k_1(u)](t,x)|^q = f_k(t,x).$$

Since $(\tilde{u}_k)$ is convergent in X, (f_k) is convergent in $L^1(Q_T)$, consequently equiintegrable in $L^1(Q_T)$, too. Hence functions $\left(a_i^{(l)}(\cdot\ , \tilde{u}_k(\cdot), D\tilde{u}_k(\cdot);u)\right)$ $(k \in \mathbb{N})$ are equiintegrable in $L^q(Q_T)$. Then by Vitali's theorem we have

$$\lim_{k\to\infty} \|a_i^{(l)}(\cdot\ , \tilde{u}_k(\cdot), D\tilde{u}_k(\cdot);u) - a_i^{(l)}(\cdot\ , u(\cdot), Du(\cdot);u)\|_{L^q(Q_T)} = 0.$$

REMARK. Observe that we have shown also the following facts: $A(\tilde{u}_k) - \tilde{A}_u(\tilde{u}_k) \to 0$ weakly in X^* and $[A(\tilde{u}_k) - \tilde{A}_u(\tilde{u}_k), v_k] \to 0$, if (v_k) is a bounded sequence in X.

COERCITIVITY. From condition F4 we get

$$[A(u),u] \ge \int_{Q_T} \left[g_2(u)|u(t,x)|^p + |Du(t,x)|^p - [k_2(u)](t,x)\right] dt\,dx =$$
$$= g_2(u)\|u\|_X^p - \|k_2(u)\|_{L^1(Q_T)},$$

thus using F4 again we obtain

$$\lim_{\|u\|_X\to\infty} \frac{[A(u),u]}{\|u\|_X} \ge \lim_{k\to\infty} \left[g_2(u)\|u\|_X^{p-1} - \frac{\|k_2(u)\|_{L^1(Q_T)}}{\|u\|_X}\right] = +\infty.$$

PSEUDOMONOTONICITY. Let us suppose that

(6) $\quad u_k \to u$ weakly in X and $D_t u_k \to D_t u$ weakly in X^*,

further

(7) $$\limsup_{k\to\infty} [A(u_k), u_k - u] \le 0.$$

By using the subsequence trick it is sufficient to show that for a subsequence $(\tilde{u}_k) \subset (u_k)$

$$\lim_{k\to\infty} [A(\tilde{u}_k), \tilde{u}_k - u] = 0 \quad \text{and} \quad A(\tilde{u}_k) \to A(u) \ \text{ weakly in } X^*.$$

Since the imbedding $W^{1,p}(\Omega) \to L^p(\Omega)$ is compact and (u_k) is bounded in X and (Du_k) is bounded in X^* by its weak convergence, hence from the well known imbedding theorem (see [4]) there exists a subsequence $(\tilde{u}_k) \subset (u_k)$ such that $\tilde{u}_k \to u$ in Y. Then by using the above remark we obtain

$$\lim_{k\to\infty} [A(\tilde{u}_k) - \tilde{A}_u(\tilde{u}_k), \tilde{u}_k - u] = 0. \tag{8}$$

Comparing this with (7) it follows that

$$\limsup_{k\to\infty}[\tilde{A}_u(\tilde{u}_k), \tilde{u}_k - u] \le 0. \tag{9}$$

We know that $\tilde{A}_u$ is pseudomonotone with respect to $D(L)$ (see [2]), hence from conditions (6) and (9) we get

$$\lim_{k\to\infty} [\tilde{A}_u(\tilde{u}_k), \tilde{u}_k - u] = 0 \quad \text{and} \quad \tilde{A}_u(\tilde{u}_k) \to \tilde{A}_u(u)(= A(u)) \text{ weakly in } X^*. \tag{10}$$

From this, by using (8) we have $\lim_{k\to\infty} [A(\tilde{u}_k), \tilde{u}_k - u] = 0$. On the other hand, we have shown in the proof of demicontinuity that $\tilde{A}_u(\tilde{u}_k) - A(\tilde{u}_k) \to 0$ weakly in X^*, so that by using the second part of (10) we obtain $A(\tilde{u}_k) \to$ $\to A(u)$ weakly in X^*. This completes the proof. ∎

COROLLARY 1. *For every $F \in X^*$ the equation*

$$D_t u + A(u) = F, \qquad u(0) = 0$$

has got a solution $u \in D(L)$.

PROOF. Since operator D_t is closed, linear and maximal monotone (see e.g. [5]), therefore the statement follows from the preceding theorem and theorem 4 in [3]. ∎

3. Examples

In this section we deal with a general form of functions $a_i^{(l)}$ which fulfil conditions F1–F5. In the end we show some concrete examples.

3.1. General case

Suppose that function $a_i^{(l)}(t,x,\zeta_0,\zeta;v)$ has the form:

$$(11)\quad a_i^{(l)}(t,x,\zeta_0,\zeta;v) = \left[H^{(l)}(v)\right](t,x)b_i^{(l)}(t,x,\zeta_0,\zeta) + \\ + \left[G^{(l)}(v)\right](t,x)d_i^{(l)}(t,x,\zeta_0,\zeta) \text{ if } i \neq 0, \text{ and}$$

$$(12)\quad a_0^{(l)}(t,x,\zeta_0,\zeta;v) = \left[H^{(l)}(v)\right](t,x)b_0^{(l)}(t,x,\zeta_0,\zeta) + \\ + \left[G_0^{(l)}(v)\right](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta),$$

where $b_i^{(l)}, d_i^{(l)}, H^{(l)}, G^{(l)}, G_0^{(l)}$ have the following properties.

K1. Functions $b_i^{(l)}\colon Q_T \times \mathbb{R}^{(n+1)N} \to \mathbb{R}$ and $d_i^{(l)}\colon Q_T \times \mathbb{R}^{(n+1)N} \to \mathbb{R}$ has the Carathéodory property. This means that they are measurable in (t,x) for every $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$, and continuous in (ζ_0,ζ) for a.e. $(t,x) \in Q_T$ $(i = 1,\dots,n;\ l = 1,\dots,N)$.

K2. There exist constants $c_1 > 0$, $0 \le r < p-1$ and a function $k_1 \in L^q(Q_T)$ such that

a) $|b_i^{(l)}(t,x,\zeta_0,\zeta)| \le c_1(|\zeta_0|^{p-1} + |\zeta|^{p-1}) + k_1(t,x)$,

b) $|d_i^{(l)}(t,x,\zeta_0,\zeta)| \le c_1(|\zeta_0|^r + |\zeta|^r)$

for a.e. $(t,x) \in Q_T$ and each $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$ $(i = 1,\dots,n;\ l = 1,\dots$ $\dots,N)$.

K3. For each $\zeta \neq \eta \in \mathbb{R}^{nN}$

a) $\displaystyle\sum_{i=1}^{n}[b_i^{(l)}(t,x,\zeta_0,\zeta) - b_i^{(l)}(t,x,\zeta_0,\eta)](\zeta_i^{(l)} - \eta_i^{(l)}) > 0$,

b) $\displaystyle\sum_{i=1}^{n}[d_i^{(l)}(t,x,\zeta_0,\zeta) - d_i^{(l)}(t,x,\zeta_0,\eta)](\zeta_i^{(l)} - \eta_i^{(l)}) \ge 0$

for a.e. $(t,x) \in Q_T$ and each $\zeta_0 \in \mathbb{R}^N$ $(l = 1,\dots,N)$.

K4. There exist a constant $c_2 > 0$ and a function $k_2 \in L^1(Q_T)$ such that

a) $\displaystyle\sum_{i=0}^{n} b_i^{(l)}(t,x,\zeta_0,\zeta)\zeta_i^{(l)} \ge c_2(|\zeta_0^{(l)}|^p + |\zeta^{(l)}|^p) - k_2(t,x)$,

b) $\sum_{i=1}^{n} d_i^{(l)}(t,x,\zeta_0,\zeta)\zeta_i^{(l)} \geq 0$

for a.e. (t,x) and each $(\zeta_0,\zeta) \in \mathbb{R}^{(n+1)N}$ $(l = 1,\dots,N)$.

K5.

a) The operator $H^{(l)}: L^p(0,T;(L^p(\Omega))^N) \to L^\infty(Q_T)$ is bounded and continuous such that for every $v \in L^p(0,T;(L^p(\Omega))^N)$ $[H^{(l)}(v)](t,x) \geq c_3 > 0$ holds for a.e. $(t,x) \in Q_T$.

b) The operators $G^{(l)}, G_0^{(l)}: L^p(0,T;(L^p(\Omega))^N) \to L^{\frac{p}{p-r-1}}(Q_T)$ are bounded, continuous where r is given in K2/b. Further, for each $v \in L^p(0,T;(L^p(\Omega))^N)$ we have $[G^{(l)}(v)](t,x) \geq 0$ for a.e. $(t,x) \in$ $\in Q_T$ and

$$\lim_{\|v\|_{L^p(0,T;V)}\to\infty} \frac{\int_{Q_T} |G_0^{(l)}(v)(t,x)|^{\frac{p}{p-r-1}}\,dt\,dx}{\|v\|^p_{L^p(0,T;V)}} = 0, \quad l = 1,\dots,N. \tag{13}$$

CLAIM 1. *Assume that conditions K1–K5 hold. Then functions defined in* (11), (12) *satisfy conditions F1–F5.*

For the proof we need a technical lemma.

LEMMA 1. *Let us introduce the following operators:*

$$[H(v)](t,x) = \sum_{l=1}^{N} |[H^{(l)}(v)](t,x)|$$

$$[G(v)](t,x) = \sum_{l=1}^{N} |[G^{(l)}(v)](t,x)|$$

$$[G_0(v)](t,x) = \sum_{l=1}^{N} |[G_0^{(l)}(v)](t,x)|.$$

Then operators H, G and G_0 fulfil the conditions formulated in K5 on $H^{(l)}$, $G^{(l)}$ and $G_0^{(l)}$, respectively.

PROOF OF LEMMA 1. We have to prove only (13) which follows easily by estimating the integrand by $|a+b|^s \leq 2^{s-1}(|a|^s + |b|^s)$. ∎

PROOF OF CLAIM 1.

CONDITION F1. From K1 obviously follows F1.

CONDITION F2. Let $i > 0$ and $r > 0$. It is obvious that

$$|[H^{(l)}(v)](t,x)b_i^{(l)}(t,x,\zeta_0,\zeta)| \leq$$
$$\leq \|H(v)\|_{L^\infty(Q_T)} \left(c_1\left(|\zeta_0|^{p-1} + |\zeta|^{p-1}\right) + k_1(t,x)\right).$$

On the other hand by using Young's inequality with conjugate exponents $1 < p_1 = \frac{p-1}{r} < \infty$ and $q_1 = \frac{p-1}{p-r-1}$ we get

$$|[G^{(l)}(v)](t,x)d_i^{(l)}(t,x,\zeta_0,\zeta)| \leq |[G(v)](t,x)d_i^{(l)}(t,x,\zeta_0,\zeta)| \leq \frac{|d_i^{(l)}(t,x,\zeta_0,\zeta)|^{p_1}}{p_1} + \frac{|[G(v)](t,x)|^{q_1}}{q_1}. \tag{14}$$

Estimating by K2/b and $|a+b|^s \leq 2^{s-1}(|a|^s + |b|^s)$ we obtain

$$|[G^{(l)}(v)](t,x)d_i^{(l)}(t,x,\zeta_0,\zeta)| \leq const \cdot \left(|\zeta_0|^{rp_1} + |\zeta|^{rp_1} + |[G(v)](t,x)|^{q_1}\right)$$
$$= const \cdot \left(|\zeta_0|^{p-1} + |\zeta|^{p-1} + |[G(v)](t,x)|^{q_1}\right). \tag{15}$$

Combining the above estimations we have

$$|a_i^{(l)}(t,x,\zeta_0,\zeta;v)| \leq const \cdot \Big[\left(\|H(v)\|_{L^\infty(Q_T)} + 1\right)\left(|\zeta_0|^{p-1} + |\zeta|^{p-1}\right) +$$
$$+ \|H(v)\|_{L^\infty(Q_T)}k_1(t,x) + |[G(v)](t,x)|^{q_1}\Big].$$

By the boundedness of operator H and by the continuous imbedding $X \to Y$ we have that $\|H(\cdot)\|_{L^\infty(Q_T)}$ is a bounded $X \to \mathbb{R}^+$ functional. Further, from $k_1 \in L^q(Q_T)$ it follows that $\|H(\cdot)\|_{L^\infty(Q_T)}k_1$ is a bounded $X \to L^q(Q_T)$ operator. Observe that $q_1 q = \frac{p}{p-r-1}$ so that

$$\int_{Q_T} \left(|[G(v)](t,x)|^{q_1}\right)^q \, dt dx = \int_{Q_T} |[G(v)](t,x)|^{\frac{p}{p-r-1}} dt dx$$
$$= \left(\|G(v)\|_{L^{\frac{p}{p-r-1}}(Q_T)}\right)^{\frac{p}{p-r-1}}. \tag{16}$$

Due to boundedness of G this means that $|G(\cdot)|^{q_1}$ is a bounded $X \to L^q(Q_T)$ operator.

Now let $r = 0$. Observe that $q_1 = 1$, moreover from K2/b we have $|d_i^{(l)}(t,x,\zeta_0,\zeta)| \le 2c_1$. So in this case we also have an inequality similar to (15):

$$|[G^{(l)}(v)](t,x)d_i^{(l)}(t,x,\zeta_0,\zeta)| \le const \cdot |[G(v)](t,x)|^{q_1}.$$

This means that this case can be treated in the same way. This completes the proof in case $i > 0$. Case $i = 0$ is the same, we only have to replace G by G_0.

CONDITION F3. Using condition K3 and K5/a we get for $\zeta \neq \eta$

$$\sum_{l=1}^{N}\sum_{i=1}^{n}\left(a_i^{(l)}(t,x,\zeta_0,\zeta;v) - a_i^{(l)}(t,x,\zeta_0,\eta;v)\right)(\zeta_i^{(l)} - \eta_i^{(l)}) =$$

$$= \sum_{l=1}^{N}[H^{(l)}(v)](t,x)\sum_{i=1}^{n}\left(b_i^{(l)}(t,x,\zeta_0,\zeta) - b_i^{(l)}(t,x,\zeta_0,\eta)\right)(\zeta_i^{(l)} - \eta_i^{(l)})+$$

$$+\sum_{l=1}^{N}[G^{(l)}(v)](t,x)\sum_{i=1}^{n}\left(d_i^{(l)}(t,x,\zeta_0,\zeta) - d_i^{(l)}(t,x,\zeta_0,\eta)\right)(\zeta_i^{(l)} - \eta_i^{(l)}) > 0.$$

CONDITION F4. Taking into account conditions K4 and K5 we obtain

$$\sum_{l=1}^{N}\sum_{i=0}^{n} a_i^{(l)}(t,x,\zeta_0,\zeta;v)\zeta_i^{(l)} = \sum_{l=1}^{N}\left[H^{(l)}(v)\right](t,x)\cdot \tag{17}$$

$$\cdot\sum_{i=0}^{n} b_i^{(l)}(t,x,\zeta_0,\zeta)\zeta_i^{(l)} + \sum_{l=1}^{N}\left[G^{(l)}(v)\right](t,x)\sum_{i=1}^{n} d_i^{(l)}(t,x,\zeta_0,\zeta)\zeta_i^{(l)}+$$

$$+\sum_{l=1}^{N}\left[G_0^{(l)}(v)\right](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\zeta_0^{(l)} \ge$$

$$\ge \sum_{l=1}^{N} c_3c_2\left(\left|\zeta_0^{(l)}\right|^p + \left|\zeta^{(l)}\right|^p\right) - c_3k_2(t,x)+$$

$$+\sum_{l=1}^{N}\left[G_0^{(l)}(v)\right](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\zeta_0^{(l)} \ge$$

$$\ge c_4c_3c_2\left(|\zeta_0|^p + |\zeta|^p\right) - c_3Nk_2(t,x)+$$

$$+\sum_{l=1}^{N}\left[G_0^{(l)}(v)\right](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\zeta_0^{(l)}.$$

In the last estimation we used inequality $|a+b|^s \le 2^{s-1}(|a|^s+|b|^s)$. Put $c' = c_4c_3c_2$ and investigate only the terms in the last sum. Let $\varepsilon > 0$ be fixed a constant such that $\frac{\varepsilon^p}{p} < \frac{c'}{3N}$, and use the ε-inequality with exponents $p,\ q$. Then we have

$$(18)\qquad |[G_0^{(l)}(v)](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\zeta_0^{(l)}| \le$$
$$\le |[G_0(v)](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\zeta_0^{(l)}| \le$$
$$\le \frac{\varepsilon^p}{p}|\zeta_0^{(l)}|^p + \frac{\varepsilon^{-q}}{q}|[G_0(v)](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)|^q.$$

The first term in the right hand side of (18) is less or equal than $\frac{c'}{3N}\left(|\zeta_0|^p + \right.$ $\left. + |\zeta|^p\right)$. In the second term using the ε-inequality with $\mu > 0$ (defined later) and exponents $p_1,\ q_1$ similarly to (14), (15), the following estimation holds:

$$(19)\quad \left|\left[G_0^{(l)}(v)\right](t,x)d_0^{(l)}(t,x,\zeta_0,\zeta)\right|^q \le$$
$$\le const\cdot\left(\mu^{p_1}\left(|\zeta_0|^{p-1}+|\zeta|^{p-1}\right)+\mu^{-q_1}|[G_0(v)](t,x)|^{q_1}\right)^q \le$$
$$\le c^*\mu^{p_1q}\left(|\zeta_0|^p+|\zeta|^p\right)+c^*\mu^{-q_1q}|[G_0(v)](t,x)|^{q_1q}.$$

Let μ be such that $\frac{c^*\mu^{p_1q}\varepsilon^{-q}}{q} < \frac{c'}{3N}$. Then substituting (18) and (19) into (17)

$$\sum_{l=1}^{N}\sum_{i=0}^{n} a_i^{(l)}(t,x,\zeta_0,\zeta;v)\zeta_i^{(l)} \ge$$
$$\ge \frac{c'}{3}\left(|\zeta_0|^p+|\zeta|^p\right) - \underbrace{(c_3Nk_2(t,x)+Nd^*|[G_0(v)](t,x)|^{q_1q})}_{=:[h(v)](t,x)}$$

where $h(v)\in L^1(Q_T)$ following from (16) (and $k_2\in L^1(Q_T)$). Moreover

$$\|h(v)\|_{L^1(Q_T)} \le c_3N\|k_2\|_{L^1(Q_T)} + Nd^*\int_{Q_T}|[G_0(v)](t,x)|^{\frac{p}{p-r-1}}dt\,dx.$$

From the lemma we know that G_0 fulfil (13), hence

$$\lim_{\|v\|_X\to\infty}\|v\|_X^{p-1}\left(\frac{c'}{3}-\frac{\|h(v)\|_{L^1(Q_T)}}{\|v\|_X^p}\right) = \lim_{\|v\|_X\to\infty}\frac{c'}{3}\|v\|_X^{p-1} = +\infty.$$

CONDITION F5. Let $r > 0$. Suppose that $u_k \to u$ weakly in X and strongly in Y. Then (u_k) is bounded in X. Therefore from K2/a follows that $\left(b_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot))\right)$ $(k \in \mathbb{N})$ is bounded in $L^q(Q_T)$, since it is easy to see (similarly to (3)) that

$$\int_{Q_T} |b_i^{(l)}(t,x,u_k(t,x), Du_k(t,x)|^q dtdx \le$$
$$\le const \cdot \int_{Q_T} \left[|u_k(t,x)|^{(p-1)q} + |Du_k(t,x)|^{(p-1)q} + |k_1(t,x)|^q\right] dtdx \le$$
$$\le const \cdot (\|u_k\|_X^p + \|k_1\|^q_{L^q(Q_T)}) \le K.$$

Further observe that $\left(d_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot))\right)$ $(k \in \mathbb{N})$ is bounded in $L^{\frac{p}{r}}(Q_T)$, since by K2/b

$$\int_{Q_T} |d_i^{(l)}(t,x,u_k(t,x), Du_k(t,x))|^{\frac{p}{r}} dtdx \le$$
$$\le \int_{Q_T} \left[|u_k(t,x)|^{r\frac{p}{r}} + |Du_k(t,x)|^{r\frac{p}{r}}\right] dtdx = \|u_k\|_X^p \le K.$$

Hence

$$\int_{Q_T} |([H^{(l)}(u_k)](t,x) - [H^{(l)}(u)](t,x))b_i^{(l)}(t,x,u_k(t,x), Du_k(t,x))|^q dtdx \le$$
$$\le \|H^{(l)}(u_k) - H^{(l)}(u)\|^q_{L^\infty(Q_T)} \int_{Q_T} |b_i^{(l)}(t,x,u_k(t,x), Du_k(t,x))|^q dtdx \le$$
$$\le K\|H^{(l)}(u_k) - H^{(l)}(u)\|_{L^\infty(Q_T)} \to 0,$$

by using the continuity of $H^{(l)}$. On the other hand, Hölder's inequality with exponents $p_1,\ q_1$ shows that

$$\int_{Q_T} |([G^{(l)}(u_k)](t,x) - [G^{(l)}(u)](t,x))d_i^{(l)}(t,x,u_k(t,x), Du_k(t,x))|^q dtdx \le$$
$$\le \left(\int_{Q_T} |d_i^{(l)}(t,x,u_k(t,x), Du_k(t,x)|^{\frac{p}{p-1}\frac{p-1}{r}} dtdx\right)^{\frac{1}{p_1}} \cdot$$
$$\cdot \left(\int_{Q_T} |[G^{(l)}(u_k)](t,x) - [G^{(l)}(u)](t,x)|^{\frac{p}{p-1}\frac{p-1}{p-r-1}} dtdx\right)^{\frac{1}{q_1}} \le$$

$$\le K^{\frac{1}{p_1}} \|G^{(l)}(u_k) - G^{(l)}(u)\|_{L^{\frac{p}{p-r-1}}(Q_T)}^{\frac{p-r-1}{p}} \to 0,$$

since $G^{(l)}$ is continuous. This means that

$$\|a_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot); u_k) - a_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot); u)\|_{L^q(Q_T)} \le \tag{20}$$
$$\le \|(H^{(l)}(u_k) - H^{(l)}(u))b_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot))\|_{L^q(Q_T)} +$$
$$+ \|(G^{(l)}(u_k) - G^{(l)}(u))d_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot))\|_{L^q(Q_T)} \to 0.$$

If $r = 0$, then the first term on the right hand side of (20) tends to 0. Since $\frac{p}{p-r-1} = q$ (hence G maps to $L^q(Q_T)$ continuously) and $|b_i^{(l)}(t, x, \zeta_0, \zeta)| \le \le 2c_1$, so that

$$\|(G^{(l)}(u_k) - G^{(l)}(u))d_i^{(l)}(\cdot\,, u_k(\cdot), Du_k(\cdot))\|_{L^q(Q_T)} \le$$
$$\le 2c_1 \|(G^{(l)}(u_k) - G^{(l)}(u))\|_{L^q(Q_T)} \to 0.$$

Hence the second term in the right hand side of (20) tends to 0, too. Case $i = 0$ can be treated similarly, replacing $G^{(l)}$ by $G_0^{(l)}$.

3.2. Concrete examples

3.2.1. Operator $H^{(l)}$

Let $\phi\colon \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\phi \ge c > 0$. Let us introduce the following operators on $L^p(0, T; (L^p(\Omega))^N)$:

$$[\tilde{H}_1(v)](t, x) := \phi\left(\int_{Q_t} \sum_{j=1}^{N} b_j v^{(j)}\right), \quad \text{where } b_j \in L^q(Q_T) \ (1 \le j \le N),$$

$$[\tilde{H}_2(v)](t, x) := \phi\left(\left[\int_{Q_t} |v|^\alpha\right]^{\frac{1}{\alpha}}\right), \quad \text{where } 1 \le \alpha \le p.$$

CLAIM 2. *The above $\tilde{H}_1$ and $\tilde{H}_2$ fulfil condition K5/a.*

PROOF. We prove only the case of $\tilde{H}_1$, the other can be made by similar techincs. From Hölder's inequality we know that $b_j v^{(j)} \in L^1(Q_T)$, so that

$\tilde{H}_1$ is well defined, and obviously $\tilde{H}_1(v) \geq c > 0$. On the other hand, if $\|v\|_Y \leq K$ then we have

$$\left| \int_{Q_t} \sum_{j=1}^N b_j v^{(j)} \right| \leq \sum_{j=1}^N \int_{Q_T} |b_j v^{(j)}| \leq K \sum_{j=1}^N \|b_j\|_{L^q(Q_T)},$$

from where by continuity of ϕ follows that $\tilde{H}_1$ maps to $L^\infty(Q_T)$ and it is bounded indeed. Further, if $(v_k) \to v$ in $L^p(0, T; (L^p(\Omega))^N)$ then we have

$$\left| \int_{Q_t} \sum_{j=1}^N b_j v_k^{(j)} - \int_{Q_t} \sum_{j=1}^N b_j v^{(j)} \right| \leq$$

$$\leq \sum_{j=1}^N \left(\int_{Q_T} |b_j|^q \right)^{\frac{1}{q}} \left(\int_{Q_T} |v_k^{(j)} - v^{(j)}|^p \right)^{\frac{1}{p}} \to 0,$$

therefore by continuity of ϕ it follows that $\tilde{H}_1(v_k) \to \tilde{H}_1(v)$ in $L^\infty(Q_T)$. This completes the proof of continuity. ∎

3.2.2. Operators $G^{(l)}$, $G_0^{(l)}$

Let $\psi: \mathbb{R} \to \mathbb{R}$ be a continuous function such that $|\psi(y)| \leq const \cdot$ $\cdot |y|^{p-r_0-1}$ holds for some $0 \leq r_0 < p - 1$. Let us introduce the following operators on $L^p(0, T; (L^p(\Omega))^N)$:

$$[\tilde{G}_1(v)](t,x) := \psi \left(\int_0^t \sum_{j=1}^N a_j(\tau,x) v^{(j)}(\tau,x) d\tau \right),$$

$$[\tilde{G}_2(v)](t,x) := \psi \left(\int_\Omega \sum_{j=1}^N a_j(t,\xi) v^{(j)}(t,\xi) d\xi \right),$$

where $a_j \in L^\infty(Q_T)$ $(1 \leq j \leq N)$,

$$[\tilde{G}_3(v)](t,x) := \psi \left(\left[\int_0^t |v(\tau,x)|^\alpha d\tau \right]^{\frac{1}{\alpha}} \right), \quad \text{where } 1 \leq \alpha \leq p.$$

CLAIM 3. *The above $\tilde{G}_i$ fulfil conditions made on $G_0^{(l)}$ in K5/b with $0 \leq r < r_0$. (If $\psi \geq 0$, then obviously the nonnegativity condition is fulfilled, too.)*

PROOF. We show only the case of operator $\tilde{G}_1$. Let be $0 \le r < r_0 < p-1$ then from properties of ψ it is obvious that

$$\int_{Q_T} |[\tilde{G}_1(v)](t,x)|^{\frac{p}{p-r-1}} dtdx \le$$

$$\le const \cdot \int_{Q_T} \left(\sum_{j=1}^{N} \int_0^T \|a_j\|_{L^\infty(Q_T)} |v^{(j)}(\tau,x)| d\tau \right)^{p\lambda} dtdx \le$$

$$\le const \cdot \int_{Q_T} \left(\cdot \sum_{j=1}^{N} \int_0^T |v(\tau,x)| d\tau \right)^{p\lambda} dtdx =$$

$$= const \cdot \int_{Q_T} \left(\int_0^T |v(\tau,x)| d\tau \right)^{p\lambda} dtdx,$$

where $0 < \lambda = \frac{p-r_0-1}{p-r-1} < 1$. By using Hölder's inequality with exponents $p_1 = \frac{1}{\lambda} (> 1)$ and $q_1 = \frac{p_1}{p_1 - 1}$ we obtain:

$$\int_{Q_T} \left(\int_0^T |v(\tau,x)| d\tau \right)^{p\lambda} dtdx \le$$

$$\le const \cdot \left(\int_{Q_T} \left(\int_0^T |v(\tau,x)| d\tau \right)^{p\lambda \frac{1}{\lambda}} dtdx \right)^{\lambda} \cdot \left(\int_{Q_T} 1^{q_1} \right)^{\frac{1}{q_1}} =$$

$$= const \cdot \left(\int_{Q_T} \left(\int_0^T |v(\tau,x)| d\tau \right)^{p} dtdx \right)^{\lambda}$$

Now we may estimate again by Hölder's inequality and after that we may use Fubini's theorem. We get

$$\int_{Q_T} \left(\int_0^T |v(\tau,x)| d\tau \right)^{p} dtdx \le$$

$$\le \int_{Q_T} \left[\left(\int_0^T |v(\tau,x)|^p d\tau \right)^{\frac{1}{p}} \left(\int_0^T 1^q d\tau \right)^{\frac{1}{q}} \right]^p dtdx =$$

$$= const \cdot \int_{Q_T} \int_0^T |v(\tau,x)|^p d\tau dx dt = const \cdot \int_{Q_T} |v(t,x)|^p dt dx \leq const \cdot \|v\|_X^p.$$

Summarizing the above estimations one gets

$$\int_{Q_T} |[\tilde{G}_1(v)](t,x)|^{\frac{p}{p-r-1}} dt dx \leq const \cdot \|v\|_X^{p\lambda}.$$

From this it is easy to see that $\tilde{G}_1$ is a bounded operator which maps to $L^{\frac{p}{p-r-1}}(Q_T)$. Further

$$\lim_{\|v\|_X \to \infty} \frac{\int_{Q_T} |[\tilde{G}_1(v)](t,x)|^{\frac{p}{p-r-1}} dt dx}{\|v\|_X^p} = \lim_{\|v\|_X \to \infty} \|v\|_X^{p(\lambda-1)} = 0,$$

since $\lambda - 1 < 0$. Continuity of the operator can be proved similarly to the previous theorem.

REMARK. From lemma it is easy too see that linear combinations of the above operators fulfil condtitions K5/a and K5/b, too.

3.2.3. Functions $b_i^{(l)}$, $d_i^{(l)}$

We show the well known examples. Let $b_i^{(l)}(t,x,\zeta_0,\zeta) := \tilde{b}_i^{(l)}(t,x,\zeta_0,\zeta_i^{(l)})$, where $\tilde{b}_i^{(l)}: Q_T \times \mathbb{R}^{N+1} \to \mathbb{R}$ is a Carathéodory function such that the following hold. Function $\zeta_i^{(l)} \mapsto \tilde{b}_i^{(l)}(t,x,\zeta_0,\zeta_i^{(l)})$ is strictly increasing,

$$\left|\tilde{b}_i^{(l)}\left(t,x,\zeta_0,\zeta_i^{(l)}\right)\right| \leq c_1 \left(|\zeta_0|^{p-1} + \left|\zeta_i^{(l)}\right|^{p-1}\right) + k_1(t,x),$$

and

$$\tilde{b}_i^{(l)}\left(t,x,\zeta_0,\zeta_i^{(l)}\right)\zeta_i^{(l)} \geq c_2 \left|\zeta_i^{(l)}\right|^p - k_2(t,x),$$

where $c_1 > 0$, $k_1 \in L^q(\Omega)$ and $k_2 \in L^1(Q_T)$. Then $b_i^{(l)}$ obviously fulfil K1, K2/a. K4/a follows by inequality $|a+b|^s \leq 2^{s-1}(|a|^s + |b|^s)$ and K3/a follows from monotonicity.

Similarly, let $d_i^{(l)}(t,x,\zeta_0,\zeta) := \tilde{d}_i^{(l)}(t,x,\zeta_0,\zeta_i^{(l)})$ $(i \neq 0)$ where $\tilde{d}_i^{(l)}: Q_T \times \mathbb{R}^{N+1} \to \mathbb{R}$ is a Carathéodory function such that the following hold. Function $\zeta_i^{(l)} \mapsto \tilde{d}_i^{(l)}(t,x,\zeta_0,\zeta_i^{(l)})$ is monotone nondecreasing, $\tilde{d}_i^{(l)}(t,x,\zeta_0,0)=0$, and

$$\left|\tilde{d}_i^{(l)}\left(t,x,\zeta_0,\zeta_i^{(l)}\right)\right| \le c_1\left(|\zeta_0|^r + \left|\zeta_i^{(l)}\right|^r\right) + k_1(t,x),$$

where $c_1 > 0$, $k_1 \in L^q(Q_T)$ and $0 \le r < p-1$. If $i=0$, then let $d_0^{(l)}$ be a Carathéodory-function which satisfies

$$\left|d_0^{(l)}(t,x,\zeta_0,\zeta)\right| \le c_1\left(|\zeta_0|^r + |\zeta|^r\right) + k_1(t,x).$$

Then conditions K1, K2/b, K3/b obviously hold. To prove K4/b we only have to observe that (if $i \neq 0$) $\tilde{d}_i^{(l)}(t,x,\zeta_0,\zeta_i^{(l)})\zeta_i^{(l)} \ge 0$.

REMARK. The simplest examples for the above general conditions are $\zeta_i^{(l)} \mapsto \zeta_i^{(l)}|\zeta_i^{(l)}|^{p-2}$ and $\zeta_i^{(l)} \mapsto \zeta_i^{(l)}|\zeta_i^{(l)}|^{r-1}$ if $r>0$. If $r=0$ let $d_i^{(l)} \equiv 0$ and $d_0^{(l)} \equiv 1$.

References

[1] L. SIMON, On parabolic functional differential equations of general divergence form, *submitted to Proceedings of FSDONA 04*.

[2] L. SIMON, On systems of strongly nonlinear parabolic functional differential equations, *Periodica Math. Hung.* **33** (1996), 145–151.

[3] J. BERKOVITS and V. MUSTONEN, Topological degree for perturbations of linear maximal monotone mappings and applications to a class of parabolic problems, *Rend. Mat. Ser.* **VII,12** (1992), Roma, 597–621.

[4] J. L. LIONS, *Quelques méthodes de résolution des problémes aux limites non linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[5] E. ZEIDLER, *Nonlinear functional analysis and its applications II A and II B*, Springer, 1990.

Ádám Besenyei
Department of Applied Analysis,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
badam@cs.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 157–185

ON GENERALIZED ROBERTSON–WALKER SPACE TIMES

By

DÁVID SZEGHY

(Received December 14, 2004)

1. Definitions, generalization of earlier statements and the concept of multivisible points

After some preliminary statements, firstly we define the multivisible points and examine the properties of these points, and then we give a description of the multivisible point structure of the warped product space-times. At the end of the section we give some examples, which show that this structure can be the same at each point of the universe, or it can vary there.

DEFINITION. Let $\mathbb{R}^1_1$ be the 1-dimensional Lorentz manifold obtained by multiplying with -1 the canonical metric of the real line, $I = (\alpha, \beta) \subset \mathbb{R}^1_1$ an open interval, P a complete 3-dimensional Riemannian manifold and $\phi : I \to \to \mathbb{R}^+$ a positive valued smooth function, then the warped product $M = I \times_\phi P$ is a 4-dimensional Lorentz manifold having a canonical time-orientation by the ordering of I and accordingly it is called a *warped product space-time*. If in particular S is a simply connected 3-dimensional Riemannian manifold of constant curvature, i.e. S is either the 3-dimensional euclidean space $\mathbb{E}^3$ or the 3-dimensional sphere $\mathbb{S}^3$ or the 3-dimensional hyperbolic space $\mathbb{H}^3$ then $M = I \times_\phi S$ is called a *Robertson–Walker space-time* ([1] pp. 129–131; [7] pp. 341–345). Furthermore, let S' be a 3-dimensional Riemannian manifold of constant curvature such that there is a locally isometric covering map

$$\omega : S \to S',$$

then the warped product $M' = I \times_\phi S'$ is called a *generalized Robertson–Walker space-time;* in this definition the case $S' = S$ is also admitted.

The following basic concepts of Lorentzian geometry will be applied subsequently: If $(M, <, >)$ is a time oriented Lorentz manifold and $p \in M$ then by the *light cone at p in* T_pM the set

$$\Lambda_p = \{v \in T_pM - \{0_p\} \mid \langle v, v\rangle = 0\}$$

is meant; moreover, by the *light cone at p in M* the image under the exponential map:

$$L_p = exp_p(\Lambda_p \cap E_pM).$$

where E_pM is the domain of the exponential map. If $(M, <, >)$ is also time oriented then the corresponding *future light cone*s Λ_p^+ in T_pM and L_p^+ in M are obviously defined.

DEFINITION. Let $(M, <, >)$ be a semi-Riemannian manifold, $\varphi : C \to M$ its geodesic where $C \subset \mathbb{R}$ is a closed interval and $\theta : J \to C$ a smooth strictly monotone increasing funtion on an interval $J \subset \mathbb{R}$. Then the smooth curve

$$\psi = \varphi \circ \theta : J \to M$$

is called a *pregeodesic* of the semi-Riemannian manifold. Let ∇ be the Levi–Cività covariant derivation of the semi-Riemannian manifold then the equation

$$\nabla_{\dot\psi}\dot\psi = \frac{d^2\theta}{d\tau^2}\dot\psi$$

is obviously satisfied by the pregeodesic ψ and conversely any smooth regular curve ψ of the semi-Riemannian manifold such that $\nabla_{\dot\psi(\tau)}\dot\psi$ is collinear with $\dot\psi(\tau)$ throughout is a pregeodesic ([7], p. 69; pp. 95–96).

By application of O'Neill's results on geodesics in warped products the following lemmas are obtainable. With the help of this lemas we will examine the lightlike geodesics in a warped product space-time. It turns out that these geodesics correspond to the geodesics of the Riemannian spacelike factor in the warped product.

LEMMA 1.1. *Let $M = I \times_\phi P$ be a warped product space-time and*

$$\zeta(\tau) = (\xi(\tau), \eta(\tau)),\ \tau \in J$$

its geodesic. Then the following are valid:

1. $J \ni \tau \mapsto \eta(\tau) \in P$ is a pregeodesic in the Riemannian manifold P.

2. *If $\kappa : \mathbb{R} \to P$ is a geodesic parametrized by arclength, then there is a unique inextendible future directed lightlike geodesic*

$$\zeta(\tau) = (\xi(\tau), \eta(\tau)), \quad \tau \in J$$

of M such that $\eta(\tau) = \kappa \circ \Theta(\tau)$, $\tau \in J$ where $\Theta : J \to \mathbb{R}$ is a strictly monotone increasing smooth function.

PROOF. By an obvious modification of the proof given for Robertson–Walker space-times (see e.g. [7], p. 353–354) where the assumption that P is of constant curvature is not essential. ∎

Next we give, in some sense, a uniform parametrization for the lightlike geodesics.

LEMMA 1.2. *Let $M = I \times_\phi P$ be a warped product space-time which is lightlike complete, where P is a complete Riemannian manifold and $p = (t, c) \in M$. Then there is a smooth function*

$$\chi_p : \mathbb{R}^+ \to I$$

such that if $\kappa : \mathbb{R} \to \{t\} \times P$ is a geodesic in the totally geodesic submanifold $\{t\} \times P$ in arclength parametrization with $\kappa(0) = p$, then

$$\mathbb{R}^+ \ni \sigma \mapsto (\chi_p(\sigma), \kappa(\sigma)) \in M$$

is a future directed lightlike pregeodesic.

PROOF. Consider a geodesic $\kappa : \mathbb{R} \to \{t\} \times P$ parametrized by arclength, and assume that there is a smooth function $\lambda : \mathbb{R}^+ \to I$ such that the curve

$$\mathbb{R}^+ \ni \sigma \mapsto (\lambda(\sigma), \kappa(\sigma))$$

is lightlike. Then

$$\begin{gathered}
0 = -\langle \dot\lambda(\sigma), \dot\lambda(\sigma) \rangle + \phi^2(\lambda(\sigma)) \langle \dot\kappa(\sigma), \dot\kappa(\sigma) \rangle, \\
(\dot\lambda(\sigma))^2 = \phi^2((\lambda(\sigma)), \\
\dot\lambda(\sigma) = \phi(\lambda(\sigma))
\end{gathered}$$

has to be valid. Then by integration

$$\int_0^{\lambda(\sigma)} \frac{d\lambda}{\phi(\lambda)} = \int_0^\sigma d\tau = \sigma$$

is obtained. But then a smooth function $\Psi(\lambda) = \sigma$, $\sigma \in \mathbb{R}^+$ is obtained and $\lambda(\sigma) = \Psi^{-1}(\sigma)$ is valid. Conversely, a smooth function of the above kind yields a lightlike curve with the given properties. The curve is a geodesic in consequence of the preceding *lemma*. ∎

Note that by "we can change the direction of the time", we mean that we can lift the above geodesic κ to a past directed lightlike pregeodesic, if we take

$$\dot{\lambda}(\sigma) = -\phi(\lambda(\sigma))$$

instead of $\dot{\lambda}(\sigma) = \phi(\lambda(\sigma))$ in the above *lemma*. So we can define χ_p on the whole real line $\mathbb{R}$ in the above *lemma*. From now on all the *lemmas* will have a dual in the above sense: "changing the direction of the time".

LEMMA 1.3. *Let $M = I \times_\phi P$ be a warped product space time were P is a complete 3-dimensional Riemannian manifold, then a lightlike geodesic intersects $\{t\} \times P$ for every $t \in I$.*

PROOF. It is enough to see that a future directed lightlike geodesic from $p = (t, c) \in M$ intersects all $\{t'\} \times P$ for every $t' > t$, $t' \in I$. Let $\zeta : \mathbb{R}^+ \to P$ be a geodesic from p parametrized by arclength and $\gamma = (\eta, \zeta)$ its unique lift to lightlike pregeodesic. For an indirect argument assume that γ is not intersecting a $\{t\}^* \times P$, $t^* > t$, $t^* \in I$. Let $\delta = \min\{\phi(s) \mid t \leq s \leq t^*\}$, then for the lightlike pregeodesic $\gamma = (\eta, \zeta) : \mathbb{R}^+ \to M$

$$0 = \langle \dot{\gamma}(s), \dot{\gamma}(s) \rangle = \phi^2(\eta(s)) \langle \dot{\zeta}(s), \dot{\zeta}(s) \rangle - \langle \dot{\eta}(s), \dot{\eta}(s) \rangle = \phi^2(\eta(s)) - \langle \dot{\eta}(s), \dot{\eta}(s) \rangle,$$
$$\phi(\eta(s)) = \dot{\eta}(s)$$

is valid. Then from $Im(\gamma) \subset [t, t^*] \times_\phi P$ and

$$t^* - t > \int_0^\infty \dot{\eta}(s) ds \geq \int_0^\infty \delta \, ds = \infty$$

a contradiction is obtained. So *lemma 1.3* follows. ∎

DEFINITION. Let $M = I \times_\phi P$ be a warped product space-time where P is a complete 3-dimensional Riemannian manifold. Fix a point $p = (t, c) \in M$, and consider a geodesic $\kappa : \mathbb{R} \to \{t\} \times P$ with $\kappa(0) = p$ and parametrized by arclength. Then in consequence of the preceding *lemmas* there is a unique smooth function

$$\chi_p : [F_p, R_p] \to I$$

such that $\sigma \mapsto (\chi_p(\sigma), \kappa(\sigma)) \in M$ is the unique lightlike pregeodesic which projects to κ and this function χ_p is called the *galactic time needed for a photon to cover the spatial distance* σ along the spacelike geodesic κ, and R_p is called the spacelike distance we can cover during the existence of the universe, correspondingly F_p is the past spacelike distance we can cover starting at the beginning of the universe.

If we see a star in the universe it means mathematically that there is a lightlike geodesics joining us and a “previous” image of the star in the space-time. Now we give the mathematical definition of the *visibility* in this sense.

DEFINITION. Let $(L, <,>)$ be a time oriented Lorentz manifold and $x, y \in \in L$. If there is a future directed lightlike geodesic $\varphi : [0,\beta] \to L$ with $\varphi(0) = x$ and $\varphi(\beta) = y$ then x is said to be *visible* from y, moreover the 1-dimensional half-space

$$T_\beta\varphi(\mathbb{R}^-) = \{\lambda \cdot \dot{\varphi}(\beta) \mid \lambda < 0,\ \lambda \in \mathbb{R}\} \subset T_y L$$

is called the corresponding *direction of visibility* of x from y; let $V(x;y)$ be the *set of directions of visibility* of x from y. If x is visible from y and there is only a single direction of visibility of x from y, in other words if $V(x;y)$ has a single element then x is said to be *simply visible* from y; if there are more than one such directions, in other words if $V(x;y)$ has more than 1 element then x is said to be *multivisible* from y. The number

$$o(x,y)$$

of the elements of $V(x,y)$ is called the *order of visibility* of x from y. If P is a 3-dimensional Riemannian manifold such that to any two different points of P there is a single geodesic joining them, then there are no multivisible points in a warped product space-time $M = I \times_\phi P$.

We remark that if we see a star from two different directions it means that there are two lightlike geodesics which have started from “previous” positions of the star and are intersecting at us “here” and “now”, but the two fotons that we see have not necessary the same “age” we mean that they could start at different galactic time, more precisely at different space-time points. In the next lemmas we will give a description for the multivisible points.

In the next lemmas we fix a point p and describe the points from which it is multivisible. We use for this description the future light cone L_p^+. But correspondingly if we want to describe the points which are multivisible from p then we should replace L_p^+ with the past light cone L_p^-. So in this sense the following results have dual roles.

LEMMA 1.4. *Let $(L, <,>), (L', <,>')$ be time-oriented Lorentz manifolds satisfying the causality condition and such that there is an isometric covering map $\omega : L \to L'$ which is also time orientation preserving, and $p, q \in \in L$, $p', q' \in L'$ such that $p' = \omega(p)$, $q' = \omega(q)$. Then p' is multivisible from q' if and only if*

$$q \in L_p^+ \cap L_{\Theta(p)}^+$$

holds for the future light cone L_p^+ with a suitable non-trivial time orientation preserving deck transformation $\Theta : L \to L$ associated with the covering map ω or p is already multivisible from q.

Note that in the time oriented Lorentz-manifold by the causality condition we do not allow a lightlike geodesic loop.

PROOF. The proof is a modification of the proof of lemma 3.1 in [8].

Let $q \in L_p^+ \cap L_{\Theta(p)}^+$, then $\tilde{q} = \Theta^{-1}(q) \in L_p^+$ is valid, since Θ is isometric therefore $\Theta(L^+{}_p) = L_{\Theta(p)}^+$. Moreover, there are future directed lightlike geodesics

$$\zeta(\sigma) = (\xi(\sigma), \eta(\sigma)),\ \sigma \in [0, \beta],\ \tilde{\zeta}(\tau) = (\tilde{\xi}(\tau), \tilde{\eta}(\tau)),\ \tau \in [0, \tilde{\beta}],$$

such that $\zeta(0) = \tilde{\zeta}(0) = p$, $\zeta(\beta) = q$, $\tilde{\zeta}(\tilde{\beta}) = \tilde{q}$. Put $p' = \omega(p)$, $q' = \omega(q)$, then the curves

$$\omega \circ \zeta : [0, \beta] \to L,\ \omega \circ \tilde{\zeta} : [0, \tilde{\beta}] \to L$$

are future directed lightlike geodesics from p' to q'. It will be shown that

$$T_\beta(\omega \circ \zeta)(\mathbb{R}^-) \neq T_{\tilde{\beta}}(\omega \circ \tilde{\zeta}(\mathbb{R}^-)$$

is valid. For sake of an indirect argument assume that the above directions are equal and consider the geodesics

$$\overleftarrow{\zeta}(\sigma) = \omega \circ \zeta(\beta - \sigma),\ \sigma \in [0, \beta],$$
$$\overleftarrow{\tilde{\zeta}}(\tau) = \omega \circ \tilde{\zeta}(\tilde{\beta} - \tau),\ \tau \in [0, \tilde{\beta}].$$

By the indirect assumption the above geodesics have the same initial point and initial direction, therefore the image of one of them is included in that of the other. But these images must be equal since otherwise there would be a lightlike geodesic loop in L' at the end point q'.

Therefore the lifts of the geodesics $\overleftarrow{\zeta}$, $\overleftarrow{\tilde{\zeta}}$ to L with the initial point q coincide, and consequently $p = \Theta(p)$ is valid. But then Θ has to be trivial in contradiction with its choice.

Now assume that there is a point $q' \in L'$ which is multivisible from the point $p' \in L'$. Then there are future directed lightlike geodesics

$$\psi : [0,\beta] \to L', \ \tilde{\psi} : [0,\tilde{\beta}] \to L'$$

such that $\psi(0) = \tilde{\psi}(0) = p'$, $\psi(\beta) = \tilde{\psi}(\tilde{\beta}) = q'$ and their tangent vectors

$$\dot{\psi}(\beta), \ \dot{\tilde{\psi}}(\tilde{\beta}) \in T_{q'}L'$$

are independent. Consider the lifts φ, $\tilde{\varphi}$ of the above geodesics in L with the initial point p and put $q = \varphi(\beta)$, $\tilde{q} = \tilde{\varphi}(\tilde{\beta})$. Then $q \neq \tilde{q}$ is valid, since otherwise p would be multivisible from q. But then there is a non-trivial deck transformation Θ associated with ω such that $q = \Theta(\tilde{q})$ holds, since $\omega(q) = \omega(\tilde{q})$ is valid. Now

$$q \in L_p^+ \cap L_{\Theta(p)}^+$$

holds by the construction above. ∎

The next definition will be very helpful to give a formal description of the multivisible points.

DEFINITION. Let P be a Riemannian manifold, and a, $b \in P$ then the set of geodesics $\gamma : [\alpha,\beta] \to P$ such that $\gamma(\alpha) = a$, $\gamma(\beta) = b$ is denoted by $\mathscr{G}(a,b)$ and called the *the system of geodesics joining* a to b; furthermore by $\mathscr{L}(a,b)$ is denoted the set of the lengths of these geodesics $\gamma \in \mathscr{G}(a,b)$ and it is called *the set of geodesic lengths* between a and b.

In the special case when $P = \mathbb{R}^3$ or $P = \mathbb{H}^3$ and $d : P \times P \to \mathbb{R}$ is the corresponding distance function, then for c, $\tilde{c} \in P$

$$\mathscr{L}(c,\tilde{c}) \cap \mathscr{L}(\Theta(c),\tilde{c}) \neq \emptyset$$

is valid if and only if $d(c,\tilde{c}) = d(\Theta(c),\tilde{c})$ holds where Θ is an arbitrary isometry.

LEMMA 1.5. *Let $M = I \times_\phi P$ be a warped product space-time where P is complete, $\omega : P \to P'$ an isometric covering map, $M' = I \times_\phi P'$ the corresponding warped product space-time, $\Theta : P \to P$ a non-trivial deck transformation and $\Theta^\times : M \to M$ the associated deck transformation. Fix $p = (t,c) \in M$ and let C be the set of those $\tilde{c} \in P$ for which $\mathscr{L}(c,\tilde{c}) \cap \cap \mathscr{L}(\Theta(c),\tilde{c}) \neq \emptyset$. Then*

$$L_p^+ \cap \Theta^\times(L_p^+) = \{(\tilde{t},\tilde{c}) \in M \mid \tilde{c} \in C, \ \tilde{t} = \chi_p(\lambda), \ \lambda \in \mathscr{L}(c,\tilde{c}) \cap \mathscr{L}(\Theta(c),\tilde{c}), \ \lambda < R_q\}$$

is valid for the intersection of the above future light cones, where R_q is the spacelike distance we can cover during the existence of the universe.

PROOF. The proof is a modification of lemma 3.3 in [8].

If $q \in L_p^+ \cap \Theta^\times(L_p^+)$ then there are two different future directed lightlike geodesics

$$\zeta : [0,1] \to M,\ \tilde{\zeta} : [0,1] \to M$$

such that the following are valid $\zeta(0) = p$, $\tilde{\zeta}(0) = \Theta^\times(p)$, $\zeta(1) = \tilde{\zeta}(1) = q$. Then their projections η, $\tilde{\eta}$ on P are pregeodesics of equal length by the *corollary* of *proposition 1.2* and this length is an element of the set

$$\mathcal{L}(c,\tilde{c}) \cap \mathcal{L}(\Theta(c),\tilde{c}).$$

Conversely, if $\lambda \in \mathcal{L}(c,\tilde{c}) \cap \mathcal{L}(\Theta(c),\tilde{c})$ then there are geodesics $\eta, \tilde{\eta} : [0,1] \to S$ of length λ with

$$\eta(0) = c,\ \tilde{\eta}(0) = \Theta(c),\ \eta(1) = \tilde{\eta}(1) = \tilde{c}.$$

But then the lift of η with initial point p and of $\tilde{\eta}$ with initial point $\Theta^\times(p)$ as lightlike geodesics, which end at the same point

$$(\chi_p(\lambda),\tilde{c}) \in L_p^+ \cap \Theta^\times(L_p^+),$$

we need here that $\lambda < R_q$. Thus the assertion of the *lemma* is established. ∎

Let $\beta(c,\Theta(c))$ denote the plane which bisects the line segment $(c,\Theta(c))$.

COROLLARY 1.1. *Let $M = I \times_\phi S$ be a Robertson–Walker space-time where $S = \mathbb{R}^3$ or $\mathbb{H}^3$ and $\omega : S \to S'$ an isometric covering map. If $p = (t,c) \in M$, then*

$$\begin{aligned} L_p^+ \cap \Theta^\times(L_p^+) &= \{(\tilde{t},\tilde{c}) \in P \mid d(c,\tilde{c}) = d(\Theta(c),\tilde{c}), \tilde{t} = \chi_p(d(c,\tilde{c})), d(c,\tilde{c}) < R_p\} \\ &= \{(\tilde{t},\tilde{c}) \in P \mid \tilde{c} \in \beta(c,\Theta(c)), \tilde{t} = \chi_p(d(c,\tilde{c})), d(c,\tilde{c}) < R_p\} \end{aligned}$$

PROOF. It is a modification of corollary 1 of [8].

In fact, if $S = \mathbb{E}^3$ or $S = \mathbb{H}^3$ then

$$\begin{aligned} &\{(\tilde{t},\tilde{c}) \in M \mid \tilde{t} = \chi_p(\lambda),\ \lambda \in \mathcal{L}(c,\tilde{c}) \cap \mathcal{L}(\Theta(c),\tilde{c})\} = \\ &= \{(\tilde{t},\tilde{c}) \in M \mid d(c,\tilde{c}) = d(\Theta(c),\tilde{c}),\ \tilde{t} = \chi_p(d(c,\tilde{c}))\} \end{aligned}$$

is valid. But then by the preceding *lemma* the assertion follows. ∎

Let $M' = I \times_\phi P'$ be a warped product space-time where P' has constant curvature and the universal covering space P of P' is $\mathbb{H}^3$ or $\mathbb{R}^3$. Furthermore let $p' = (t', c') \in M'$ be a point and $\omega : P \to P', \Gamma, \Theta, \Theta^\times$ as above. Then it is easy to see, that in this case in *lemma 1.4* there is no multivisible points in $M = I \times_\phi P$. And we mentioned preceding *lemma 1.4* that *lemma 1.4, lemma 1.5* and *corollary 1.1* have dual forms. So if $p \in M$ such that $\omega(p) = p'$ then we must know only the set $L_p^- \cap \Theta^\times(L_p^-)$ for every non-trivial deck transformation $\Theta^\times$. But from the dual of *corollary 1.1* we know that $L_p^- \cap \Theta^\times(L_p^-) = \{(\tilde{t}_{\tilde{c}}, \tilde{c}) \mid \tilde{c} \in \beta(c, \Theta(c)),\ \tilde{t}_{\tilde{c}} = \chi_p(-d(c, \tilde{c})),\ d(c, \tilde{c}) < F_p\}$. We can say that by the projection $\Pi : I \times_\phi P' \to P'$ we lifted $\beta(c, \Theta(c))$ with suitable galactic time coordinates, see *figure 1*. Doing the above lifting for all bisectors, for all $\beta(c, \Theta(c))$, $\Theta \in \Gamma - \{Id\}$ we get all the points which are multivisible from p'.

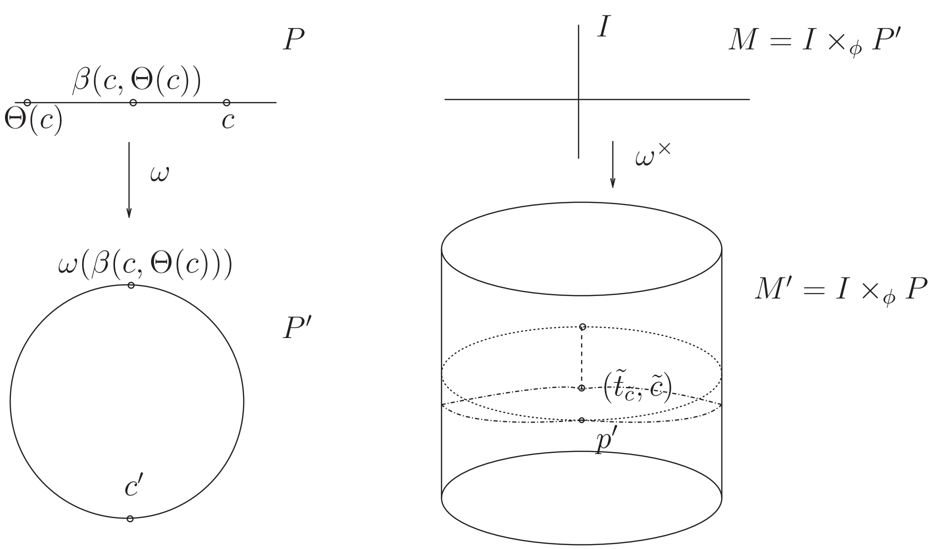


Figure 1. Getting the multivisible points, 2 dimensions of P and P' are suppressed

COROLLARY 1.2. *Let $M = I \times_\phi S$ be a Robertson–Walker space-time where $S = \mathbb{S}^3$ and $\omega : S \to S'$ an isometric covering map. If $p = (t, c) \in M$ then the following holds:*

$$L_p^+ \cap \Theta^\times(L_p^+) = \{(\tilde{t}, \tilde{c}) \mid d(c, \tilde{c}) = d(\Theta(c), \tilde{c}),\ \tilde{t} = \chi_p(|d(c, \tilde{c}) + 2k\pi|),$$
$$k \in \mathbb{Z},\ |d(c, \tilde{c}) + 2k\pi| < R_p\} =$$

$$= \{(\tilde{t},\tilde{c}) \mid \tilde{c} \in \beta(c,\Theta(c)),\ \tilde{t} = \chi_p(|d(c,\tilde{c})+2k\pi|),\ k \in \mathbb{Z},\ |d(c,\tilde{c})+2k\pi| < R_p\}$$

where $\Theta^\times$ is a deck transformation induced by a non-trivial deck transformation Θ associated with with the isometric covering map $\omega : S \to S'$. Moreover, the set of those points $q = (\tilde{t},\tilde{c}) \in M$ from which $p = (t,c)$ is multivisible is given by

$$\{(\chi_p(k\pi), A^k(c)) \mid k \in \mathbb{N} - \{0\},\ k\pi < R_p\}$$

where $A : \mathbb{S}^3 \to \mathbb{S}^3$ is the antipodal map.

Proof. It is a modification of corollary 2 of [8].

In fact

$$\{(\tilde{t},\tilde{c}) \in M \mid \tilde{t} = \chi_p(\lambda),\ \lambda \in \mathcal{L}(c,\tilde{c}) \cap \mathcal{L}(\Theta(c),\tilde{c})\} =$$
$$= \{(\tilde{t},\tilde{c}) \in M \mid d(c,\tilde{c}) = d(\Theta(c),\tilde{c}),\ \tilde{t} = \chi_p(|d(c,\tilde{c}) + 2k\pi|),\ k \in \mathbb{Z}\}$$

is valid. The second assertion of the *corollary* is obviously valid, because the geodesics from $c \in \mathbb{S}^3$ can intersect only at $A(c)$ or c. ∎

Lemma 1.6. *Let $(L,<,>)$ be a space-time satisfying the causality condition, $M = I \times_\phi P$ a warped product space-time and $\omega : M \to L$ a locally isometric time orientation preserving covering map. If $p = (t,c) \in M$, then $p' = \omega(p)$ is multivisible from a $q' = \omega(q)$, $q \in M$ if and only if one of the following two conditions is satisfied:*

1. *p is already multivisible from q*
2. *Let $p^* = (\tilde{t},\tilde{c}) \in \omega^{-1}(p')$ and $\{\mathcal{L}(\tilde{c},c') + \chi_p^{-1}(\tilde{t})\} = \{\ d + \chi_p^{-1}(\tilde{t}) \mid d \in \in \mathcal{L}(\tilde{c},c')\}$, where if $\tilde{t} < t$, then $\chi_p^{-1}(\tilde{t})$ means $-\chi_{p^*}^{-1}(t)$. Then $q \in \in \{(t',c') \mid \exists \lambda \in \mathcal{L}(c,c') \cap \{\mathcal{L}(\tilde{c},c') + \chi_p^{-1}(\tilde{t})\},\ t' = \chi_p(\lambda),\ \lambda < R_p\}$ for some $p^* = (\tilde{t},\tilde{c}) \in \omega^{-1}(p')$.*

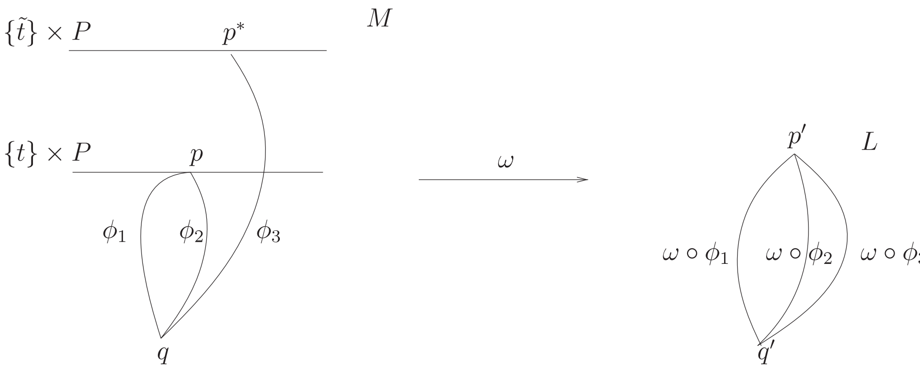


Figure 2. The multivisible points are of two kinds

PROOF. From *lemma 1.4* follows that those points, from which p' is multivisible, are of two kinds. A point of the first kind is the projection to L of a point $x \in M$ for which there is a point in $\omega^{-1}(p')$ which is multivisible from x (the first case in the *lemma*). We can get a point of the second kind if we fix a $p = (t, c) \in \omega^{-1}(p')$ and we project the intersection $L_p^+ \cap L_{p^*}^+$ to L for every $p^* = (\tilde{t}, \tilde{c}) \in \omega^{-1}(p')$, $p^* \neq p$. It can be shown as in *lemma 1.4* that for the description of ${L^+}_p \cap {L^+}_{p^*}$ condition 2 above is valid, because let $q = (t', c') \in L_p^+ \cap L_{p^*}^+$ then there are lightlike geodesics $\gamma : [0, \beta] \to M, \tilde{\gamma} : [0, \tilde{\beta}] \to M$ for which $\gamma(0) = p, \tilde{\gamma}(0) = p^*, \gamma(\beta) = \tilde{\gamma}(\tilde{\beta}) = $ $= q$. Then let their projections on P be ζ, $\tilde{\zeta}$, for which $\zeta(0) = c$, $\tilde{\zeta}(0) = \tilde{c}$, $\zeta(\beta) = \tilde{\zeta}(\tilde{\beta}) = c'$ is true. The lenght of ζ is $\chi_p(t')$ and of $\tilde{\zeta}$ is $\chi_{p^*}(t')$. From $\chi_p(\tilde{t}) + \chi_{p^*}(t') = \chi_p(t')$ (because of $p^* = (\tilde{t}, \tilde{c})$) follows that ζ is longer than $\tilde{\zeta}$ by $\chi_p(\tilde{t})$. So

$$ {L^+}_p \cap {L^+}_{p^*} \subset \{(t', c') \mid \exists\lambda \in \mathscr{L}(c, c') \cap \{\mathscr{L}(\tilde{c}, c') + \chi_p^{-1}(\tilde{t})\},\ t' = \chi_p(\lambda),\ \lambda < R_p\} $$

is valid, and for the inverse including let $\lambda \in \mathscr{L}(c, c') \cap \{\mathscr{L}(\tilde{c}, c') + \chi_p^{-1}(\tilde{t})\}$ then there are geodesics $\zeta : [0, \beta] \to P, \tilde{\zeta} : [0, \tilde{\beta}] \to P$ for which $\zeta(0) = c$, $\tilde{\zeta}(0) = \tilde{c}$, $\zeta(\beta) = \tilde{\zeta}(\tilde{\beta}) = c'$ hold where ζ is longer than $\tilde{\zeta}$ by $\chi_p^{-1}(\tilde{t})$, so their lifts to lightlike pregeodesics end in the same point $(\chi_p(\lambda), c')$. ∎

By the next theorem we will show that, if the spacelike factor in a warped product space-time has some good properties, then there exist points in "good position" in the universe; namely these are the points with the minimal galactic time coordinate from which we see a point of the surface of last scatter in opposite directions. There are yet points in "bad position"; namely, these are those which have the graetest galctic time coordinate for which we see a point from the last scattering surface also from different but not necessary in opposite directions; namely if (t, c) is a point in "bad position" then for no $t' < t$ galactic time value can we see from (t', c) a point of the surface of last scatter in different directions. We give some examples too where we examine how the above points can occure.

THEOREM 1.1. *Let P be a simply connected complete Riemannian manifold where to any two points there is only one geodesic line passing through them, $\omega : P \to P'$ a locally isometric covering map such that P' is compact and let $M' = I \times_{\phi} P'$ be a corresponding warped product space-time.*

Fix a $t \in I$ and consider the set Ξ of those $\tilde{t} \in I$, $t < \tilde{t}$ for which there are pairs $(c, \tilde{c}), c, \tilde{c} \in P'$ such that

$$o(p', q') \geq 2$$

where $p' = (t, c)$, $q' = (\tilde{t}, \tilde{c}) \in M'$. Consider also the set Υ of those $\tilde{t} \in I$ for which there is a $c \in P'$ such that for any (t', c') with $c' \in P'$ and $t \leq t' \leq \tilde{t}$ the following holds:

$$o(p', q') \leq 1$$

where $p' = (t, c)$, $q' = (t', c') \in M'$. Put now

$$t_* = \inf\{\tilde{t} \mid \tilde{t} \in \Xi\},$$
$$t^* = \sup\{\ \tilde{t} \mid \tilde{t} \in \Upsilon\ \}.$$

Then there is a closed geodesic γ_* in P' such that

1_*. *The length of γ_* is equal to $2\chi_{p'}^{-1}(t_*)$.*

2_*. *γ_* is not homotopic to 0.*

3_*. *γ_* is the shortest one among those closed curves in P' which are not homotopic to 0.*

Moreover, there is a $c \in P'$ such that

1^*. *$o((t, c), (t', c')) \leq 1$ holds for any $(t', c') \in P'$ with $t \leq t' < t^*$.*

2^*. *There is a geodesic loop in P' of length $2\chi_{p'}^{-1}(t^*)$, with a possible corner at c, and not homotopic to 0.*

PROOF. Note that Ξ or Υ may be empty and then the geodesics given in the theorem do not exist.

Fix an arbitrary point $c \in P'$ and a $\hat{c} \in P$ such that $c = \omega(\hat{c})$ holds. Put

$$\varrho(c) = \frac{1}{2}\min\{d(\hat{c}, \Theta(\hat{c})) \mid \Theta \in \Gamma - \{I\}\}$$

where Γ is the deck transformation group associated with ω and I is its identity element. In fact, $\varrho(c)$ does not depend on the choice of $\hat{c}$ in P.

It will be shown now that the function $c \mapsto \varrho(c)$, $c \in P'$ is continuous. It is enough to show that each $c_0 \in P'$ has such a neighbourhood that ϱ is continuous on it. Since ω is a covering map, there is a neighbourhood U of c_0 such that

$$\omega^{-1}(U) = \cup\{\tilde{U}_i \mid i \in \mathbb{N}\}$$

and $\omega\lceil\tilde{U}_i$, $i \in \mathbb{N}$ is a diffeomorphism onto U. Fix one $\tilde{U} = \tilde{U}_1$ of these neighbourhoods. Then $\sigma = \omega^{-1}\lceil U : U \to \tilde{U}$ is a diffeomorphism and the choice $\hat{c} = \sigma(c)$, $c \in U$ can be made continuously. Therefore the function

$$U \ni c \mapsto d(\hat{c}, \Theta(\hat{c})) = d(\sigma(c), \Theta(\sigma(c))),\ c \in U$$

is continuous for each fixed $\Theta \in \Gamma - \{I\}$ by the continuity of the distance function d (see e.g. [5], pp. 156–159). Since the action of the group Γ on P is properly discontinuous, there is a neighbourhood $W \subset U$ of c_0 such that the sets $\Theta(W)$, $\Theta \in \Gamma$ are pairwise disjoint. Therefore there is a finite number Θ_i, $i = 1, \ldots, k$ of deck transformations such that

$$\begin{aligned}\varrho(c) &= \inf\{d(\sigma(c), \Theta(\sigma(c))) \mid \Theta \in \Gamma - \{I\}\}\\ &= \inf\{d(\sigma(c), \Theta_i(\sigma(c))) \mid i = 1, \ldots, k\},\quad c \in W\end{aligned}$$

is valid. But the infimum of a finite number of continuous functions is continuous.

Now let $c_* \in P'$ be a point where ϱ attains its minimum. Fix a $\hat{c}_* \in P$ with $c_* = \omega(\hat{c}_*)$; then there is a geodesic $\hat{\gamma}_*$ of length $2\varrho(c_*)$ from $\hat{c}_*$ to $\Theta(\hat{c}_*)$ with some $\Theta \in \Gamma$. Therefore

$$\gamma_* = \omega \circ \hat{\gamma}_*$$

is a geodesic loop at c_* and γ_* has minimal length among all geodesic loops in P'. In fact, if $\gamma : [0,\beta] \to P'$ with $\gamma(0) = \gamma(\beta) = x$ is a geodesic loop at x, then fix an $\hat{x} \in \omega^{-1}(x)$ and consider the lift $\hat{\gamma} : [0,\beta] \to P$ of γ with $\hat{\gamma}(0) = \hat{x}$ and put $\hat{x}' = \hat{\gamma}(\beta)$. Then there is a $\Xi \in \Gamma - \{I\}$ with $\hat{x}' = \Xi(\hat{x})$. But then

$$\mathscr{L}(\gamma) = \mathscr{L}(\hat{\gamma}) = d(\hat{x}, \hat{x}') \geq 2\varrho(x) \geq 2\varrho(c_*) = \mathscr{L}(\gamma_*)$$

is valid. In order to prove by contradiction that γ_* cannot have a corner at c_* consider the arclength parametrization $\gamma_* : [0, \beta_*] \to P'$ and assume that γ_* has a corner at c_*, i.e.

$$\dot{\gamma}_*(0) \neq \dot{\gamma}_*(\beta)$$

holds. Then $\Theta(\dot{\hat{\gamma}}_*(0)) \neq \dot{\hat{\gamma}}(\beta)$. Put now $f = \hat{\gamma}_*(\frac{\beta}{2})$, then

$$d(f, \Theta(f)) < d(\hat{c}_*, \Theta(\hat{c}_*))$$

is valid and yields a contradiction. The assertion that γ_* is not homotopic to 0 is an obvious consequence of the construction.

Now let $c^* \in P'$ be a point where ϱ attains its maximum. Fix a point $\hat{c}^* \in \omega^{-1}(c^*)$, then there is a $\Theta \in \Gamma - \{I\}$ with

$$\varrho(c^*) = \frac{1}{2} d(\hat{c}^*, \Theta(\hat{c}^*))$$

and if $\hat{\gamma}^* : [0, \beta^*] \to P$ is a geodesic from $\hat{c}^*$ to $\Theta(\hat{c}^*)$ in arclength parametrization then $\omega \circ \hat{\gamma}^*$ is a geodesic loop at c^* which is not homotopic to 0. Let now $\zeta = (\xi, \eta)$, $\tilde{\zeta} = (\tilde{\xi}, \tilde{\eta})$ be the future directed lightlike geodesics in M' starting from (t, c^*) such that η, $\tilde{\eta}$ are those pregeodesics which in arclength parametrization are equal respectively to $\gamma^* \lceil [0, \frac{1}{2}\beta^*]$, $\overleftarrow{\gamma}^* \lceil [\beta^*, \frac{1}{2}\beta^*]$. Then ζ, $\tilde{\zeta}$ have a common endpoint (t^e, c^e). In order to show by contradiction that $t^e \in \Upsilon$ is valid assume that there is a point $(t', c') \in M'$ such that $t' < t^e$ and

$$\# V((t, c^*), (t', c')) > 1$$

holds for the number of the elements of $V((t, c^*), (t', c'))$. Therefore there are two different geodesics in P' from c^* to c'. Therefore $\omega^{-1}(c')$ cannot intersect the interior of the fundamental domain F of c^* which is the generalized Dirichlet cell obtained by the *"méthode de rayonnement"* of É. Cartan ([1], pp. 183–186). But $t' < t^e$ implies that $d(c^*, c') < \frac{1}{2} d(c^*, \Theta(c^*))$ is valid and therefore $\omega^{-1}(c')$ has to be in the fundamental domain F. In order to show that $t^e = t^*$ holds observe that the existence of a $\tilde{t} \in \Upsilon$ with $t^e < \tilde{t} \leq t^*$ contradicts the defintion of t^e. ∎

In the proof of *theorem 1.1* we fixed an arbitrary point $c \in P'$ and a $\hat{c} \in P$ such that $c = \omega(\hat{c})$ holds and put

$$\varrho(c) = \frac{1}{2} \min\{d(\hat{c}, \Theta(\hat{c})) \mid \Theta \in \Gamma - \{I\}\}$$

where Γ is the deck transformation group associated with ω and I its identity element. In fact, we have shown in the proof that the $\varrho(c)$ does not depended on the choice of $\hat{c}$ in P, and the minimum of this continuous function $\varrho : P' \to \mathbb{R}^+$ was t_* and its maximum was t^*.

Now we show some properties of ϱ and give some examples for 3-dimensional P.

LEMMA 1.7. *Let P a complete Riemannian manifold, $\omega : P \to P'$ a locally isometric covering map, Γ the deck transformation group associated with ω and*

$$\varrho(c) = \frac{1}{2} \inf\{d(\hat{c}, \Theta(\hat{c})) \mid \Theta \in \Gamma - \{I\}\},$$

where $\hat{c} \in P$, $\omega(\hat{c}) = c$. If the function $\varrho : P' \to \mathbb{R}^+$ is constant with value r in a neighbourhood of $x \in P'$, then there is a $\Theta_j \in \Gamma$ and a geodesic γ parametrized by arclength with $Im(\gamma) \subset P$ and its neighbourhood V, such that $\Theta_j(\gamma(t)) = \gamma(t + 2r)$ and if $\hat{y} \in V$ and ζ is the geodesic parametrized by arclength joining $\hat{y}$ and $\Theta_j(\hat{y})$, then $\Theta_j(\zeta(t)) = \zeta(t + 2r)$. So we can say that Θ_j is "like a translation" along γ.

PROOF. The function ϱ is continuous by the proof of the above *theorem*. Moreover, W, in the proof of the above *theorem* can be taken as $W = \sigma(B(x, \delta))$, where σ is defined in the proof of the above *theorem* and $B(x, \delta)$ is the open ball. Furthermore we can assume that δ is so small that the function ϱ has a constant value r on $\omega(W) = B(x, \delta)$. As in the proof of the above *theorem* we can choose deck transformations $\Theta_1, \ldots, \Theta_k \in \Gamma - \{Id\}$ such that $\varrho(c) = \frac{1}{2} \min\{d(\hat{c}, \Theta_i(\hat{c})) \mid i = 1, \ldots, k,\ \omega(\hat{c}) = c, \}$ if $c \in \omega(W)$. Put

$$G_i = \{y \in \omega(W) \mid d(\sigma(y), \Theta_i(\sigma(y))) \neq 2r\},\ F_i = W - G_i$$

where G_i are open sets. Then there is a $j \in \{1, \ldots, k\}$ for which $int F_j \neq \emptyset$. To prove the last statement make the indirect assumption that $int F_i = \emptyset$, $\forall i \in \{1, \ldots, k\}$. The sets G_i are dense in $\omega(W)$, so $G_1 \cap \ldots \cap G_k \neq \emptyset$ but $\forall y \in \omega(W)$, $\exists j \in \{1, \ldots, k\}$, $d(\sigma(y), \Theta_j(\sigma(y))) = 2r$ yield $y \notin G_j$, and this contradicts the indirect assumption. So let $y \in int F_j$, $\epsilon > 0$ for which $B(y, \epsilon) \subset F_j$. Let $\gamma : \mathbb{R} \to P$ be the arclenght parametrized geodesic for which $\gamma(0) = \sigma(y)$, $\gamma(2r) = \Theta_j(\sigma(y))$. Then $\Theta_j(\gamma(\epsilon)) = \gamma(2r + \epsilon)$ is valid, because

$$2r = d(\gamma(\epsilon), \Theta_j(\gamma(\epsilon))) \leq d(\gamma(\epsilon), \gamma(2r)) + d(\gamma(2r), \Theta_j(\gamma(\epsilon))) = 2r - \epsilon + \epsilon$$

follows from the arclenght parametriztion of γ and from $d(\gamma(2r), \Theta_j(\gamma(\epsilon))) = d(\Theta_j(\gamma(0)), \Theta_j(\gamma(\epsilon))) = d(\gamma(0), \gamma(\epsilon)) = \epsilon$ by the isometry of Θ_j. But in the above inequality equality sign holds if and only if there is no break at $\gamma(2r)$, i.e. $\Theta_j(\gamma(\epsilon))$ is on γ. So from $\Theta_j(\gamma(0)) = \gamma(2r)$, $\Theta_j(\gamma(\epsilon)) = \gamma(2r+\epsilon)$ the *lemma* follows because we can repeat this construction for every $z \in B(y, \epsilon)$ and for the geodesic between $\sigma(z)$, $\Theta_j(\sigma(z))$ which will form the neighbourhood V of $Im(\gamma)$. ∎

COROLLARY 1.3. *If the function ϱ has a constant value r at $x \in P'$, then there exists a $\Theta \in \Gamma$ and $\hat{x}$, where $\omega(\hat{x}) = x$ for which Θ is translating the arclength parametrized geodesic between $\hat{x}$, $\Theta(\hat{x})$ with parameter value $2r$.*

PROOF. Consider the preceding *lemma* in the case $\mathbb{R}^+ \ni \delta \to 0$; then there is a sequence $\hat{y}_n \to \hat{x}$, Θ_{jn} for which the preceding *lemma* is true, where $j_n \in \{1, \dots, k\}$. We can suppose that $\Theta_{jn} = \Theta_j$, $\forall n \in \mathbb{N}$ but then the geodesics between $\hat{y}_n$, $\Theta_j(\hat{y}_n)$ are converging in each compact interval in $\mathbb{R}$ to the geodesic between $\hat{x}$, $\Theta_j(\hat{x})$. and because the converging sequence of geodesics is invariant under Θ (modulo a translation with $2r$ parameter value) so the *corollary* is proved. ∎

COROLLARY 1.4. *If $P = \mathbb{H}^3$ then the function ϱ can not be locally constant.*

PROOF. By the preceding *lemma* there is a $B(\hat{y}, \epsilon)$ and Θ_j with the property given there, and let $\gamma : \mathbb{R} \to \mathbb{H}^3$ denote the arclenght parametrized geodesic between $\hat{y}$, $\Theta_j(\hat{y})$. Let H_1 be the hyperplane orthogonal to γ at $\hat{y}$, it is mapped by Θ_j to H_2, the hyperplane orthogonal to γ at $\Theta_j(\hat{y})$, so that the points near to $\hat{y}$ in H_1 have distance $2r$ from their images in H_2. But by such a transformation $d(H_1, H_2)$ can attain its minimum only between $\hat{y}$, $\Theta_j(\hat{y})$. ∎

EXAMPLE 1.1. Let $P = \mathbb{H}^3$ then in case of a locally isometric covering $\omega : P \to P'$ the function ϱ can not be locally constant. So if P' is compact, $t^* \neq t_*$ holds in *theorem 1.1*.

EXAMPLE 1.2. Let $P = \mathbb{R}^3$, and let Γ be the deck transformation group generated by 3 independent translations, then the corresponding function ϱ is constant. So $t^* = t_*$ in *theorem 1.1*.

EXAMPLE 1.3. Let $P = \mathbb{R}^3$ and e_1, e_2, e_3 an ortonormated basis. Let Γ be generated by:

$$\Theta_1 : \mathbb{R}^3 \ni a \to a + 100e_2,\ \Theta_2 : \mathbb{R}^3 \ni a \to a + 100e_3$$

and by Θ_3, which is the product of the translation with e_1 alog the e_1 axis and of the rotation by π around this axis. Then it is easy to show that in the point $z = (0, 0, 25)$ the function ϱ has locally the constant value 2, because from the Dirichlet cell sructure

$$d(z, \Theta(z)) \geq 50,\ \Theta \in \Gamma,\ \Theta \neq \Theta_3^{2k},\ k \in \mathbb{Z}$$

$$\Theta_3^{2k}(a) = a + 2ke_1,$$

where we mean by generalized Dirichlet cell, the one obtained by the "méthode de rayonnement" of É. Cartan ([1], pp. 183–186).

As we saw, if the universal covering space of the spacelike component is hyperbolic, then there must be points from which the universe looks different, the multivisible point srtucture varies, the models are not homogenous; in the euclidian case there can be modells which are homogenous in this sense. For a possible model in the spherical case see [6].

2. Reconstructing the topology

Next we will show, how can we reconstruct the topology of the universe with the method given in [9], [3], assuming that in the warped product spacetime $M' = I \times_{\phi} P'$ the factor P' has constant curvature. For a pair $p' \in M'$ and $q' = (t', c') \in M' = I \times_{\phi} P'$ the directions of visibility $V(p', q')$ are in $T_{q'}M$, but in practice we measure their projections on $T_{q'}P'$, the spacelike factor of $T_q M$; so by considering these directions we will mean, that we consider their projections. If $\omega : P \to P'$ is a locally isometric covering where $P = \mathbb{H}^3$, $\mathbb{E}^3$ or $\mathbb{S}^3$ and Γ is the deck transformation group associated with ω, and if we know P and Γ or a Dirichlet cell given by Γ and its side pairing, then we can get P' uniquely. Now we will show how to construct a Dirichlet cell and Γ by means of the background radition. Mathematically this means that we fix a $\{t\} \times P'$, the surface of last scatter, and we examine the multivisible points of this section from a $q' \in M'$ far enough from the surface of last scatter.

Definition. Fix a $\{t\} \times P' \subset I \times_{\Phi} P' = M'$ and a point $q' = (t', c') \in M'$ such that $t < t'$. Let $G \subset \{t\} \times P'$ denote the set of points which are multivisible from q'. Then

$$\nu(t, q') := \bigcup \{ V(p', q') \mid p' \in G \}.$$

Note that the set $V(p', q') \subset T_{c'}P'$ is the union of half-lines from the origin.

PROPOSITION 2.1. *Let $M = I \times_{\phi} P$ be a warped product space-time, where $P = \mathbb{R}^3$ or $\mathbb{H}^3$, and $\omega : P \to P'$ an isometric covering map, such that P' is compact, furthermore fix a $\{t\} \times P$, $t \in I$. Then there exists a $t^{\diamond} \in I$, $t < t^{\diamond}$, such that if the set $\nu(t, q)$ is given for fixed $t' > t^{\diamond}$, $t' \in I$ and $q \in \{t'\} \times P'$, then the Dirichlet cell and the decktransformation group Γ corresponding to ω can be constructed, provided that we know the curvature of P and R_p is great enough.*

PROOF. Fix a point $x \in P$, and let D_x denote the Dirichlet cell of x defined by the deck transformation group Γ associated with ω. Our goal is to show, how we can reconstruct D_x and its side pairing. By *corollary 1.1* the projection of the set of those points on P' which are multivisible from $\omega((t', x))$ is obtainable as follows: Put

$$\omega(\beta(x, \Theta(x))) = \{\omega(y) \mid y \in P,\ d(x, y) = d(\Theta(x), y)\},$$

where $\Theta \in \Gamma - \{Id\}$ and then the above set of point is

$$\bigcup \{\omega(\beta(x, \Theta(x))) \mid \Theta \in \Gamma - \{Id\}\}.$$

But from the already mentioned dual form of *lemma 1.2* the equality

$$L^{-}_{(t', x)} \cap (\{t\} \times P) = \{t\} \times \partial B(x, \chi_t^{-1}(t'))$$

follows. So those points in $\{t\} \times P'$ which are multivisible from an $\omega((t', x))$ are points of the set

$$\bigcup_{\Theta \in \Gamma - \{Id\}} \{t\} \times \omega(\partial B(x, \chi_t^{-1}(t')) \cap H_{\Theta}) =$$
$$= \bigcup_{\Theta \in \Gamma - \{Id\}} \{t\} \times \omega(\partial B(x, \chi_t^{-1}(t')) \cap \partial B(\Theta(x), \chi_t^{-1}(t'))),$$

which follows from the dual of *lemma 1.4* and *1.5*, (see also *figure 3, 4*), and if $\chi_t^{-1}(t')$ is great enough then $\partial B(x, \chi_t^{-1}(t'))$ intersects all the planes defining the sides of D_x.

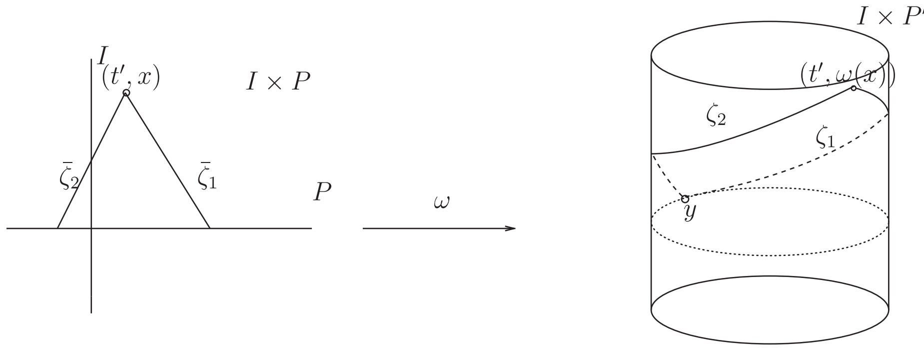


Figure 3. 2 dimensions of P are suppressed

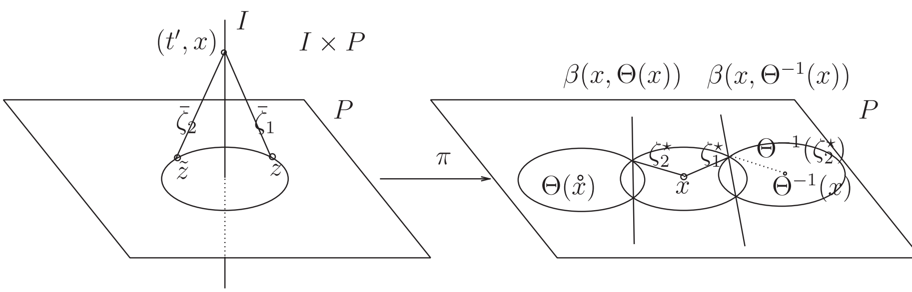


Figure 4. 1 dimension of P is suppressed

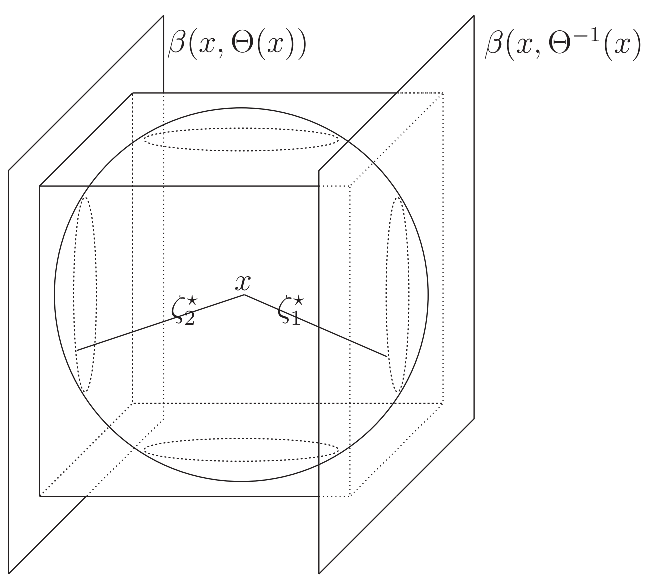


Figure 5. The multivisible points are drawing out the Dirichlet cell

By the compactness of P' there is a radius ϵ such that $B(x,\epsilon)$ intersects all the hyperplanes including the sides of the Dirichlet cell D_x for all $x \in P$. We will show that every $t^\diamond \geq \chi_t(\epsilon)$ is a good choice. Let $t' > t^\diamond$ and $\epsilon' := = \chi_t^{-1}(t')$ and $S(x,\epsilon') := \partial B(x,\epsilon')$. Fix the point $(t',\omega(x)) \in \{t'\} \times P'$ and let $y \in \{t\} \times P'$ be multivisible from $(t',\omega(x))$ and ζ_1, ζ_2 two different lightlike goedesics joining them, in fact there is at least two such lightlike joining geodesics. Let $\bar{\zeta_1}$, $\bar{\zeta_2}$ be their lifts ending at (t',x), where the beginning points are $z, \tilde{z} \in \{t\} \times P$, and $\Theta^\times$ the isometry for which $\Theta^\times(\tilde{z}) = z$. Let $\zeta_1^\bigstar$, $\zeta_2^\bigstar$ be their projections on P. These are half-lines from x. Let us repeat the above construction for all possible $y \in \{t\} \times P'$ which is multivisible from $(t',\omega(x))$ and for all possible lightlike joining geodesics. Then we get a set of halflines from x in P, which set we can get the following way: consider the set $\nu(t,(t',\omega(x)))$. After the canonical identification of $T_{(t',\omega(x))}P'$ with $T_{(t',x)}P$ and of $T_{(t',x)}P$ with P by the exponential map where the origin goes to x, we obtain the above set of halflines from x. We mean that if we carry out the method desrcibed in *figure 3* for all $y \in \{t\} \times P$ which is multivisible from $(t',\omega(x))$ and for all $\zeta_1, \zeta_2, \ldots$ lightlike joining geodesics, then the halflines $\zeta_1^\bigstar, \zeta_2^\bigstar, \ldots$ from x are the halflines in $\nu(t,(t',\omega(x)))$. The union of these half-lines intersects the sphere $S(x,\epsilon')$ in some sets and these sets are equal to the intersection of $S(x,\epsilon')$ with the planes $\beta(x,\Theta(x)) = \{y \mid d(x,y) = = d(\Theta(x),y)\}$, $\Theta \in \Gamma - \{I\}$; in fact, we have seen in *figure 3* that the end points of $\zeta_1^\bigstar, \zeta_2^\bigstar, \ldots$, which are not equal to the point x, are the projections of the multivisible points in $\{t\} \times P$ onto P, which is the union of the sets $\partial B(x,\chi_t^{-1}(t')) \cap \beta(x,\Theta(x))$ for $\Theta \in \Gamma - \{Id\}$ as we saw at the begining of the proof. The above sets can be decomposed in circle pairs, since from

$$\Theta^{-1}(\beta(x,\Theta(x)) \cap S(x,\epsilon')) = \beta(x,\Theta^{-1}(x)) \cap S(x,\epsilon')$$

we get the pairing $\beta(x,\Theta(x)) \cap S(x,\epsilon') \leftrightarrow \beta(x,\Theta^{-1}(x)) \cap S(x,\epsilon')$. This means that we see an immersed circle $S^1 \subset P'$ from two different direction. This pairing can be given by the halfline pairing. If we would know ϵ' then the circle pairs on $S(x,\epsilon')$ would give planes in pairs $\beta(x,\Theta(x))$, $\beta(x,\Theta^{-1}(x))$ by the following construction: the S^1 pairs are defined as the intersections of $S(x,\epsilon')$ with some planes; it is easy to see that the union of the halflines from x through these points, is the set $\nu(t,(t',\omega(x)))$, which is a union of some rotationally symmetric cones. If we take a cone C which is rotation symmetric around its axis E and a point $v \in S(x,\epsilon') \cap C$ then there is an unique line F

through v which intersects E orthogonally. Let us take the plane H defined by the rotation of F around E, then $H \cap S(x, \epsilon') = S(x, \epsilon') \cap C$. So we can get pairs of planes $\beta(x, \Theta(x))$, $\beta(x, \Theta^{-1}(x))$, which will define the Dirichlet cell D_x and the plane pairs which are defining D_x, there can be planes which are not of this kind, will define the side pairing transformations,which are isometries both in the euclidian case and in the hyperbolic one. Because if C, C' is the cone pair and E, E' are their axes and F, F' the corresponding lines which are defining the planes H, H', then put $m = F \cap E$, $m' = F' \cap F'$ $v, w \in S(x, \epsilon') \cap C$, furthermore let $v', w' \in S(x, \epsilon') \cap C'$ be the corresponding multivisible points. Then there is an unique side pairing transformation which takes $\Theta(x\overline{m}) \subset E'$, $\Theta(m) = m'$, $\Theta(x) \neq x$, $\Theta(v) = v'$, $\Theta(w) = w'$. By gluing together the Dirichlet cells along an edge, as desribed in *Poinceré's polyhedron theorem* the sum of the angles at the edge must be exactly 2π. In the euclidian case the homothety has no effect on the angles, so if we take $\epsilon' = 1$ we get D_x and Γ.

In the hyperbolic case if we take the sphere $S(x, 1)$ and intersect it with the directions of visibility we will get S^1 circle pairs, because the visibility directions form cones and each of them is rotation symmetric around its axis. As we saw each S^1 is an intersection of the $S(x, 1)$ with a plane, but these planes do not form a Dirichlet-cell but some of these form a cell around x. And this holds if we take $S(x, \alpha)$ instead of $S(x, 1)$, furthermore if we apply a homothety of ratio $\alpha > 1$ the angles of the planes in the cell around x will be less and this depends continuously and in a stricly monoton way on α, so there will be exactly one value α such that by glueing together the cells, with the side pairing transformations, along an edge the sum of the angles at the edge will be exactly 2π; in fact, by the 3-dimensional Caley-Klein model and by the construction of the hyperplanes from the S^1 circles as described above it is easy to show that for all value α the cells around x have always corresponding sides which yield esentially the same structure, and that there is a single good value of α which yields a Dirichlet cell and no other value of α will be good. So in the hyperbolic case we will get the radius R' moreover the Dirichlet cell and the side pairing transformations. ∎

In the $P = \mathbb{S}^3$ case we need a particular observation.

LEMMA 2.1. *Let* $P = \mathbb{S}^3$, *and* $\omega : P \to P'$ *a locally isometric covering map, and* Γ *the deck transformation group associated with* ω, *such that* $\Gamma \neq \langle A \rangle$ *i.e.* Γ *is not the group generated by the antipodal map* $A : \mathbb{S}^3 \to \mathbb{S}^3$.

Then in the canonical construction of the Dirichlet cell D_x the antipodal map A does not play role for $x \in \mathbb{S}^3$.

PROOF. Let $\Theta \in \Gamma$, $\Theta \neq Id$ and $\beta(x, \Theta(x)) = \{y \in \mathbb{R}^4 \mid d(x,y) = = d(\Theta(x), y)\}$. Then $\beta(x, \Theta(x))$ cuts the $\mathbb{S}^3 \subset \mathbb{R}^4$ in two half balls. Let γ be a geodesic through x, it is a 2π long closed geodesic, and $\beta(x, \Theta(x))$ cuts this into two geodesic arcs of length π. Let D_x denote the Dirichlet cell of x by Γ in $\mathbb{S}^3$, and for an indirect argument assume that A yields a side pair in D_x, let these be denoted by F_A, $F_{A^{-1}}$ and let $y \in intF_A$ and γ a geodesic through x and y. Then $A(\gamma) = \gamma \ni A(y) \in F_{A^{-1}}$ is valid and γ is a 2π long closed geodesic which has got a piece of length π in D_x between y and $A(y)$ and none of the $\beta(x, \Theta(x))$ planes intersects this piece. But we can take a longer piece from this geodesic as π, which has no point of intersection with $\beta(x, \Theta(x))$, $\Theta \in \Gamma - \{A, Id\}$, because y and $A(y)$ are interior points of their sides. But $\Gamma - \{A, Id\} \neq \emptyset$, and by the first remark at the beginning at the proof we saw such a geodesic arc through x can not exist. ∎

PROPOSITION 2.2. *Let $M = I \times_\phi P$, $P = \mathbb{S}^3$, and $\omega : P \to P'$ an isometric covering map, Γ the deck transformation group associated with ω and $M' = I \times_\phi P'$ the corresponding warped product space-time, where $\{t\} \times P'$ is the surface of last scatter. Assume that $\Gamma \neq \langle A \rangle$ i.e. Γ is not the group generated by the antipodal map $A : \mathbb{S}^3 \to \mathbb{S}^3$ and consider the function*

$$\mathbb{S}^3 \ni x \mapsto R_x = max \left\{ \frac{1}{2} d(x, \Theta(x)) \mid \Theta \in \Gamma - \{A, Id\} \right\} < \frac{\pi}{2},$$

Fix an $x \in \mathbb{S}^3$ then for $t' > t$, $t' \in I$, such that

$$R_x < \chi_t^{-1}(t') - \left[\frac{\chi_t^{-1}(t')}{\pi} \right] \cdot \pi < \pi - R_x .$$

If $v(t, (t', x))$ is given, then we can reconstruct the Dirichlet cell of x.

PROOF. Let $x \in \mathbb{S}^3$ and $B(x, R) \subset \mathbb{S}^3$, where $R_x < R < \pi - R_x$, mod π then

$$\beta(x, \Theta(x)) \cap B(x, R) \neq \emptyset, \quad \Theta \in \Gamma - \{A, Id\},$$
$$\beta(x, \Theta(x)) = \{y \in \mathbb{R}^4 \mid d(x, y) = d(\Theta(x), y)\}$$

is valid, and the intersection is an S^1 and we can apply the method used in the previous *proposition* as in the case of $\mathbb{H}^3$, because by the previous *lemma* $\beta(x, A(x))$ gives no side of the Dirichlet cell. ∎

So if we have the matching circle pairs from which we can reconstruct the topology of the universe, then we can tell, what curvature the universe has in the sense, that if we carry out the construction of *proposition 2.1* for the euclidian case and the sum of the angles at the equivalent edges is less than 2π then $P = \mathbb{S}^3$, if the sum is equal to 2π then $P = \mathbb{R}^3$ and if greater than 2π then $P = \mathbb{H}^3$. This can be seen by the effect of the homothety on the angles. More information about the research on the possible curvature of the universe can be found in [4].

Now a simple generalization of *proposition 2.1* is given.

PROPOSITION 2.3. *Let $M = I \times_\phi P$, $M' = I \times_\phi P'$ be warped products, such that in P there is no geodesic loop and for every two points there is a single geodesic joining them, $\omega : P \to P'$ a locally isometric covering map where P' is compact. Fix a $\{t\} \times P'$, $t \in I$ and a $c' \in P'$, then there are values $t_{c'}, t^{c'} \in I$ with $t < t_{c'} < t^{c'}$ such that by system of sets $\{\nu(t, (t', c')) \subset T_{c'}P' \mid t_{c'} < t' < t^{c'}\}$ the Dirichlet cells in P can be reconstructed, provided that $R_{(t,c')}$ is great enough.*

PROOF. Let $c \in P$, $\omega(c) = c'$ we will show (1) the map exp_c^{-1} from P to T_cP is a homeomorphism, so if we take the Dirichlet cell (D_c) in P we can take its image $(\hat{D}_c)$ in T_cP by exp_c^{-1}. If we would know this image and the side pairing, then the factorspace, factorizing by the side pairing, would be homeomorphic to P'. In (2) we will show that we can construct a homeomorphic image, to the side pairing, of $\hat{D}_c$.

(1) By the compactness of P' and the *Hopf–Rinow theorem* we get the completness of P' and since ω is locally isometric we get the geodesic completness of P. Moreover exp_c^{-1} is a homeomorphism, because exp_c is smooth and surjective; it is injective, because there is a unique geodesic parametrized by arclength between two point in P.

(2) Let $S(c, r)$ be the sphere of radius r around c, there is a maximal r_c such that $S(c, r) \cap D_c = \emptyset$ for $r < r_c$, and there is a minimal r^c such that

$B(c, r)$ contains the D_c Dirichlet cell if $r > r^c$. By the property of exp_c the following holds

$$exp_c^{-1}(S(c, r)) = S(\vec{0}, r) \subset T_c P$$

and

$$exp_c^{-1}(S(c, r) \cap D_c) = exp_c^{-1}(D_c) \cap S(\vec{0}, r).$$

But we can get $exp_c^{-1}(D_c) \cap S(\vec{0}, r)$ as the intersection of $S(\vec{0}, r)$ with the multivisible directions,

$$\nu(t, (\chi_t(r), c')) :=$$
$$:= \bigcup \{ V(y, (\chi_t(r), c')) \mid y \in \{t'\} \times P',\ y \text{ is multivisible from } (\chi_t(r), c')\},$$

which is a set of halflines from $\vec{0} \in T_c P$, as in *proposition 2.1* (see *figure 4*).

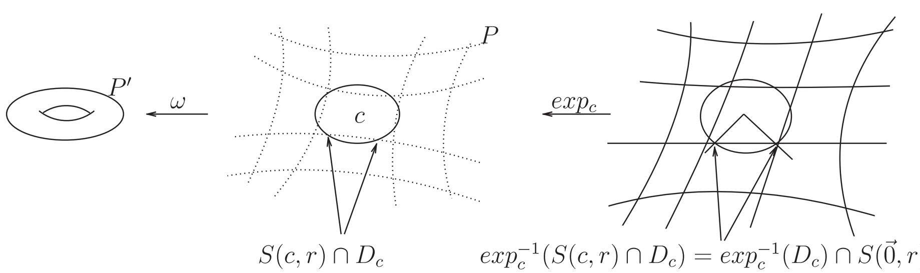


Figure 6. Lifting the Dirichlet cell

If we would know the function $\chi_t^{-1}(t') = r$ we could get the set $\hat{D}_c \subset \subset T_c P$ from the sets $\nu(t, (\chi_t(r), c'))$ for every $t_{c'} = \chi_t(r_c) < t' < t^{c'} = \chi_t(r^c)$. We mean, we could get $exp_c^{-1}(D_c) \cap S(\vec{0}, \chi_t^{-1}(t'))$ for every $t_{c'} < t' < < t^{c'}$ and the union of these would "wipe out" $\hat{D}_c$ on $T_c P$, $S(\vec{0}, \chi_t^{-1}(t')) \cap \cap (\bigcup_{\Theta \in \Gamma - \{Id\}} exp_c^{-1}(\beta(x, \Theta(x))))$ which is the intersection of $S(\vec{0}, \chi_t^{-1}(t'))$ with the multivisible half-lines from $\vec{0}$, $S(\vec{0}, \chi_t^{-1}(t')) \cap \nu(t, \chi_t(r), c')$, (see *figure 5*), and we would get the side pairing. The above method wipes out the sets $exp_c^{-1}(\beta(x, \Theta(x)))$ and the cell which arose, with $\vec{0}$ in it, will be the $\hat{D}_c$ cell.

Note that r^c must be in the domain of χ_t this is the condition that $R_{(t,c')}$ is great enough.

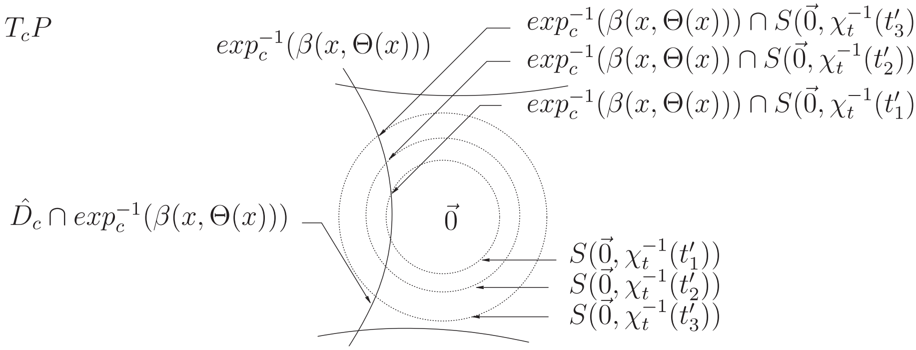


Figure 7. The circles wipe out the wall of the cell $\hat{D}_C$ on $T_C P$ for $t_C < t_1' < t_2' < t_3' < t^c$

But we don't know the function χ^{-1}, we know only that χ^{-1} is a stricly monotone increasing function. Let $\varrho : T_C P \to T_C P$ be the homeomorphism which is fixing the origin and if γ^+ is a ray from the origin parametrized by arclength then $\varrho : \gamma^+(\chi_t^{-1}(t')) \to \gamma^+(t')$ for all $t' > t$, $t' \ni I$, where we note that $\chi_t^{-1} : [t,\infty] \cap I \to [0, R_{(t,c')}]$ is a stricly monotone increasnig bijective function. Then

$$\begin{aligned} \varrho \circ exp_c^{-1}(S(c,\chi_t^{-1}(t')) \cap D_c) &= \varrho(exp_c^{-1}(D_c) \cap S(\vec{0},\chi_t^{-1}(t'))) \\ &= \varrho \circ exp_c^{-1}(D_c) \cap S(\vec{0},t') \end{aligned}$$

is valid. And we can get the set $\varrho \circ exp_c^{-1}(D_c) \cap S(\vec{0},t')$ as the intersection of $S(\vec{0},t')$ with the multivisible directions $\nu(t,(t',c'))$ from $\vec{0} \in T_C P$. So we can get $\varrho \circ exp_c^{-1}(D_c)$ with the "side pairing", which after the factorization is homeomorphic with P'. ∎

Taking an other aspect we said in *proposition 2.1, 2.2* that if we know exp_c, or selfintersections of its image e.g. the duble points, the circles, on the intersection of Λ_c, the light cone, with a suitable spacelike hyperplane, we can reconstruct the warped product M'. In a little more general case when M' is a warped product as in *proposition 2.3*, if we know exp_c, or its selfintersections, the duble points, on the intersection of the light cone with $\vec{v} \cdot [a,b] + T_C P$ we can reconstruct the topology of M', where $\vec{v}$ is the timelike vector orthogonal to $T_C P$ and + is the Minkowski addition and $[a,b]$ is a suitable interval.

Now we mention a simple generalization of *proposition 2.3*. The last *proposition* said that if we do not see any circels on the sky, and the universe is such as in the *proposition*, then we can obtain the topology of the universe if we have enough time. Now we prove that if there are circles on the sky we can get the topology, if we have enough galactic time.

LEMMA 2.2. *Let P be a 3-dimensional Riemannian manifold such that there is no geodeic loop and for every two point in P there is a unique geodesic, parametrized by arclenght, joining them. If c_1, c_2 are two distinct points. Then*

$$H := \{x \in P \mid d(c_1, x) = d(c_2, x)\}$$

is a smooth embedded submanifold, homeomorphic with $\mathbb{R}^2$.

PROOF. It can be prooved by standart methods. ∎

REMARK. We said in the preceding *proposition* that with the help of the visibility directions we can obtain the S^1 circle pairs on $S^2 = S(\vec{0}, 1) \subset T_c P$. It can be seen from the above *lemma*, that for a circle pair which devides S^2 into two pairs of spherical caps, we can choose one cap from each of these two pairs, continuously in the galactic time, given by the parameter r, such that if there is a point x in one of the chosen caps then it will stay there if the time, and consequently the value r, is increasing.

DEFININTION. If there are ω, P, P' as above, where P is the universal covering space, we call an open set $F \subset P$ a *fundamental domain* if $\omega(\bar{F}) = P'$ and $\omega|_F$ is injective.

THEOREM 2.1. *Let $M = I \times_\phi P$, $M' = I \times_\phi P'$ be warped products, such that in P there is no geodesic loop and for every two points there is a single geodesic joining them, and $\omega : P \to P'$ a locally isometric covering map such that P' is compact, and Γ the deck transformation group associated with ω. Fix a $\{t\} \times P'$, $t \in I$ and a $c' \in P'$ then for every $t^\star > t$ there are $t_{c'}$, $t^{c'}$ with $t^\star < t_{c'} < t^{c'}$, $t_{c'}, t^{c'} \in I$ such that alone by means of the multivisibility directions $\nu(t, (t', c'))$ from $\vec{0} \in T_c P$ for every $t_{c'} < t' < t^{c'}$, $t' \in I$, we can construct a topological space homeomorphic with P', provided that $R_{(t,c')}$ is great enough.*

PROOF. As in *proposition 2.3* we shall construct a ball which we shall increase as there. If we see the matching circle pairs on the sky we will see "new born circles" also, by the compactness of P' it will happen in some time. If we take such a circle pair in S^2 it is first two points and then it grows to an S^1 pair which cuts out two pairs of spherical caps and from these we can choose one by one such that by increasing the radius r, the time parameter t' is increasing; these will be the two caps which grow, by the remarks after the last *lemma*. To this circle pair belongs a pair $\beta(x,\Theta(x))$, $\beta(x,\Theta^{-1}(x))$ of "hyperplanes", as in *proposition 2.3*, such that $S(c,r) \cap (\beta(x,\Theta(x)) \cup \cup \beta(x,\Theta^{-1}(x)))$ is our circle pair. But if we take the ball $B(c,r)$ around c then it can be easily shown with the previous *lemma* that the "hyperplanes" $\beta(x,\Theta(x)), \beta(x,\Theta^{-1}(x))$ cut out two "half balls" corresponding to the above caps, topologically D^3 balls, from $B(c,r)$. By ω the images of these two balls glue together to a ball, B_r^{ω}, in P', but this is not a smooth immersed D^3 it can have a break at an S^1. Note that if we take the two "half balls" and glue these together with the transformation Θ, we get a ball, D_r^{Θ}, in P, which is smooth except a break along an embedded circle S^1 in $\beta(x,\Theta(x))$. And this is a ball, which is increasing, being blown up such as in *proposition 2.3*, because $\beta(x,\Theta(x))$ cuts the space P into two parts, in the first part are the points nearer to c and in the second the points nearer to $\Theta(c)$. The two glued half balls are in distinct parts and each is an increasing, growing, "half ball" suitable parts of the growing $B(c,r)$, which are embedded in P and stricly growing, we mean that if there is a point in this "half ball" then during the "growing" it will remain in the "half ball"; and the circle $S^1 \subset \beta(x,\Theta(x))$, which is on both "half balls", is embedded, by the above *lemma*. So the glued ball D_r^{Θ} is an embedded "ball" which is stricly growing. But topologically it is the same case as in *proposition 2.3*, because if we take this growing ball D_r^{Θ} and its images by Γ, then we see the same. Embedded balls are being stricly growing. If we take one of these balls and the points in P which are first in this ball, under the above simultaneous and uniform growing, then these will give a fundamental domain from which if we factorize with the corresponding points at the boundary, which go to the same point under the map ω, then we get a space homeomorphic to P'. So the method is the next: we will take a ball being growing in the euclidean space $B(\vec{0},r)$ we consider it as the homeomorphic immage of D_r^{Θ} where the two "half balls" are $\mathbb{R}^3_+ \cap B(\vec{0},r)$ and $\mathbb{R}^3_- \cap B(\vec{0},r)$. Let us consider $S(\vec{0},r) = exp_c^{-1}(S(c,r)) \subset T_cP$; on this sphere we have the two, growing, caps which are diffeomorphic images of the two caps of ∂D_r^{Θ}. So if we would have the corresponding self intersection points of $\omega(\partial D_r^{\Theta})$ on these two caps in $S(\vec{0},r) \subset T_cP$, then taking a homeomorphism with $\mathbb{R}^3_+ \cap S(\vec{0}r) \subset \mathbb{E}^3$ and $\mathbb{R}^3 \cap S(\vec{0},r) \subset \mathbb{E}^3$, continuous in r, then we would get a homeomorphism between D_r^Θ and $B(\vec{0},r) \subset \mathbb{E}^3$ with the corresponding self intersection points of $\omega\,(\partial D_r^\Theta)$, which can be derived from the intersection points of $\bar{\Theta}(D_r^\Theta) \cap D_r^\Theta$, $\bar{\Theta} \in \Gamma -$ $- \{Id\}$. So for this construction we must only know the corresponding self intersection points of $\omega\,(\partial D_r^\Theta)$ on this two caps in $S(\vec{0},r) \subset T_c P$, which can be constructed from the multivisibility directions (see *figure 8*).

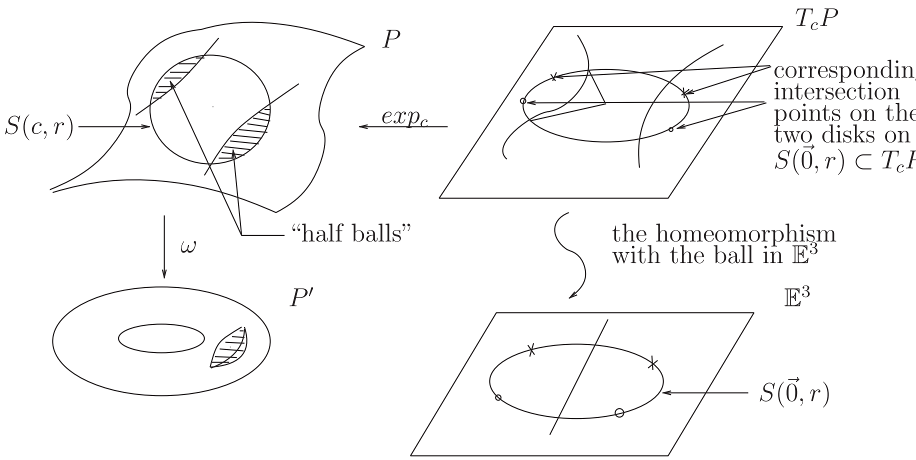


Figure 8. The construction

If there is a selfintersection point of $\omega\,(D_r^\Theta)$ then this lifts to two points y_1, y_2 in $S(c,r)$, which are also in the union of the two D^2 caps. But if two points have the same image by ω then there is a transformation $\tilde{\Theta}$ under which one of these two points goes to the other, let us assume $\tilde{\Theta}(y_1) = y_2$. So y_2 is in $S(c,r)$ and $y_2 = \tilde{\Theta}(y_1)$ is in $\tilde{\Theta}(S(c,r)) = S(\tilde{\Theta}(c),r)$ therefore it is in $S(c,r) \cap S(\tilde{\Theta}(c),r) \subset \beta(c,\tilde{\Theta}(c))$ and $\tilde{\Theta}^{-1}(y_2) = \tilde{\Theta}^{-1}(\tilde{\Theta}(y_1)) = y_1$. But this shows that the two points are in the $\beta(c,\tilde{\Theta}(c)), \beta(c,\tilde{\Theta}^{-1}(c))$ hyperplanes and in the union of the D_2 caps, which are in $S(c,r)$, so the two points are in corresponding S^1 circle pairs which are both intersecting the union of the two D^2 caps and these corresponding points are in the union of the two D^2 caps. And if two coherent points, for a S^1 pair, are in the union of the two D^2 caps, then they go by ω to the same point in P'. So we got the description of the selfintersection of S_r^ω; these are the images of the, from circle pairs derived, point pairs for which the two points are in the union of the caps. So if we apply the method of *proposition 2.3* then we get a fundamental domain on $T_c P$ with side pairings which is homeomorphic to a fundamental domain on P.

ACKNOWLEDGEMENTS. My thanks are due to J. Szenthe for his helpful remarks.

References

[1] H. BUSEMANN, *The Geometry of Geodesics*, New York, 1995.

[2] N. J. CORNISH and J. R. WEEKS, Measuring the Shape of the universe, *Notices of the AMS* **45** (1998), 1463–1471.

[3] N. J. CORNISH D. N. SPERGEL and G. D. STARKMAN, Circles in the sky; Finding topology with microwave background radiation, *Class. Quant. Grav.* **15** (1998), 2657–2670.

[4] G. F. R. ELLIS, The shape of the Universe, *Nature* **425** (2003), 566–567.

[5] D. GROMOLL, W. KLINGENBERG and W. MAYER, *Riemannsche Geometrie in Grossen*, Lecture Notes in Mathematics, **55**, 1968, Berlin.

[6] J. P. LUMINET, J. R. WEEKS, A. RIAZUELO, R. LEHOUCQ and J. P. UZAN, Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmical microwave background, *Nature* **425** (2003), 593–595.

[7] B. O'NEILL, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, 1983.

[8] D. SZEGHY and J. SZENTHE, On generalized Robertson–Walker space-times, *Proceedings of the International Conference in Geometry and Topology*, Cluj-Napoca, 2002.

[9] J. R. WEEKS, Reconstructing the global topology of the universe from the cosmic microwave background radiation, *Class. Quant. Grav.* **15** (1998), 2599–2604.

Dávid Szeghy

Department of Geometry,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
Pf.: 120
1518 Budapest, Hungary
szeghy@cs.elte.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 187–205

GENERALIZED RULED SURFACES IN THE k-ISOTROPIC n-SPACE I_n^k

By

ŽELJKA MILIN ŠIPUŠ

(Received February 4, 2002; Revised December 17, 2004)

1. Introduction

A ruled surface in the 3-dimensional Euclidean space E_3 is a surface swept out by a straight line moving along a curve. It can be parametrized by

$$\mathbf{x}(t, v) = \mathbf{c}(t) + v\,\mathbf{e}(t), \quad t \in I \subset \mathbf{R}, \ v \in \mathbf{R},$$

where c is a base curve or directix and the lines determined by $\mathbf{e}(t)$, $t \in I$, are the generators or the rulings of a given ruled surface. By the condition that the normal at infinity of the ruling $\mathbf{e}(t)$ is orthogonal to the surface normal, we get a point on a ruling $\mathbf{e}(t)$ which is a striction point of the ruling. All striction points describe the striction curve of the surface.

A ruled surface in a simply isotropic space I_3^1 and doubly isotropic space I_3^2 is defined in the same way ([1], [6], [7]). The notion of the striction point on a ruling is transfered to these spaces as a point on a ruling in which the tangent plane and the asymptotic plane (i.e. the tangent plane of a ruling at infinity) are orthogonal. In terms of isotropic spaces this happens on the admissible surfaces (whose tangent planes are non-isotropic almost in all points) in a point of a ruling in which the tangent plane is isotropic and the asymptotic plane non-isotropic. According to this definition it is shown in [1], [6], [7] that in the spaces I_3^1 and I_3^2 there are ruled surface with striction curve (whose position can vary with respect to the absolute figure of the spaces) and without it (conoidal surfaces with isotropic or absolute line as a directrix at infinity).

Mathematics Subject Classification (2000): 53A35

In this paper we develop the theory of generalized ruled surfaces in n-dimensional k-isotropic space I_n^k. The space I_n^k is introduced in [9] and studied in [4].

As in the papers [2], [3], where the same problem is treated for the n-dimensional Euclidean space E_n, these surfaces are formed by an one-parameter family of m-dimensional subspaces of I_n^k.

The classification of the generalized ruled surfaces in I_n^k on the surfaces for which the asymptotic and tangent bundle coincide (generalization of tangent surfaces) and on the surfaces for which these bundles differ in dimension by 1 (generalization of skew surfaces) is the same as in E_n. The differences in I_n^k appear in regarding the striction space of skew surfaces. In E_n in the points of the striction space the tangent $(m+1)$-plane and the asymptotic bundle are orthogonal. By defining the striction space of a ruled $(m+1)$-surface in I_n^k in this way, we demand that in these points the tangent $(m+1)$-planes are isotropic while the asymptotic bundle is non-isotropic. Since this is not always the case, in I_n^k there exist skew surfaces for which the striction space exists and for which it does not exist. Among the latter ones there exist conoidal surfaces whose generators are parallel to a certain isotropic plane.

For skew ruled surfaces in I_n^k we introduce the ith parameter of distribution which has the analogous geometrical interpretation (in terms of isotropic angles) as the ith parameter of distribution of skew ruled surfaces in E_n. It is a generalization of the parameter of distribution of ruled surfaces in E_3 given by the formula

$$p = \frac{(\dot{\mathbf{c}}, \mathbf{e}, \dot{\mathbf{e}})}{\dot{\mathbf{e}}^2} = \frac{\sigma}{\kappa},$$

where σ is the striction (described as the angle between the tangent vector of the striction curve and the generator) of a ruled surface, κ its curvature.

Finally, we apply obtained results on the 2-dimensional ruled surfaces and ruled hypersurfaces in I_n^k. However, for the latter ones, it is shown additionally that among non-cylindrical skew surfaces there exist only surfaces with striction space and conoidal surfaces with the generators parallel to some isotropic plane.

The space I_n^k is a pair (A, V) where A is a real n-dimensional affine space of points and V its corresponding vector space of translations decomposed in a direct sum of subspaces

$$V = U_1 \oplus U_2$$

satisfying $\dim U_1 = n - k$, $\dim U_2 = k$. By $B_2 = \{\mathbf{b}_{n-k+1}, \dots, \mathbf{b}_n\}$ a basis for the subspace U_2 is denoted. In U_2 a flag of vector spaces $U_2 := C_1 \supset \supset \dots \supset C_l \supset C_{l+1} \supset \dots \supset C_k := [\mathbf{b}_n]$, $C_l = [\mathbf{b}_{n-k+l}, \dots \mathbf{b}_n]$ is defined and fixed. According to it we distinguish the following classes of vectors: the Euclidean vectors as the vectors in $V \setminus U_2$ and the isotropic vectors of degree l or l-isotropic vectors, $l = 1, \dots k$, as the vectors in C_l.

The space U_1 is endowed with a Euclidean scalar product $\langle\ ,\ \rangle : U_1 \times \times U_1 \to \mathbf{R}$ which is extended on the whole V by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \pi_1(\mathbf{x}), \pi_1(\mathbf{y}) \rangle,$$

where $\pi_1 : V \to U_1$ denotes the canonical projection. In this way a semi-definite scalar product on V is defined.

Since the isotropic length $||\mathbf{x}|| := |\pi_1(\mathbf{x})|$ of an l-isotropic vector $\mathbf{x}$ is 0 we define the lth- range of $\mathbf{x}$ as

$$[\mathbf{x}]_l := x^{n-k+l}, \quad l = 1, \dots, k,$$

where x^{n-k+l} denotes the $\mathbf{b}_{n-k+l}$-coordinate of $\mathbf{x}$.

2. The Natural Basis

Let $I \subseteq \mathbf{R}$ be an open interval, O a fixed origin in I_n^k and $c : I \to I_n^k$ a C^1-curve given by its position vector $t \mapsto \mathbf{c}(t)$. The theory of curves in I_n^k is studied in [5]. Let on I be given an one-parameter family of m-dimensional subspaces $E_m(t) \subset I_n^k$, $t \in I$, $1 \le m \le n-2$.

In the subspaces $E_m(t)$ let us define an orthonormal basis $\{\mathbf{e}_1(t), \dots, \mathbf{e}_m(t)\}$ such that

$$\mathbf{e}_i : I \to V, \quad i = 1, \dots m,$$

are functions of class C^1 on I. Let us assume that, if $m + 1 \le n - k$, the vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ are all Euclidean, and if $m + 1 = n - k + r$, for some $r \in \{1, \dots, k-1\}$, that they are either Euclidean or isotropic vectors of degree $\in \{1, \dots, r\}$ such that each isotropic degree appears at most once. More precisely, in the last case there are either $n - k$ Euclidean vectors in the basis of $E_m(t)$ and $r - 1$ isotropic vectors of degree $1, \dots, i-1, i+1, \dots, r$, or $n - k - 1$ Euclidean vectors and r isotropic vectors of degree $1, \dots, r$. Every isotropic degree appears only once.

DEFINITION 1. The set of points $U \subset I_n^k$ given by its parametrized position vectors

$$\mathbf{x}(t; v_1, \ldots, v_m) = \mathbf{c}(t) + \sum_{i=1}^{m} v^i \mathbf{e}_i(t), \quad t \in I; v_1, \ldots, v_m \in \mathbf{R}$$

such that

$$\mathrm{Rank}(\dot{\mathbf{c}} + \sum_{i=1}^{m} v^i \dot{\mathbf{e}}_i; \mathbf{e}_1, \ldots, \mathbf{e}_m) = m + 1$$

defines a regular $(m+1)$-dimensional C^1-surface in I_n^k. Such a surface is called an $(m+1)$-ruled surface. The curve c is called the base curve or the directrix of U and each subspace $E_m(t)$ is called the generating or ruling space of U.

The conditions on the rulings $\mathbf{e}_1, \ldots, \mathbf{e}_m$ mentioned before are necessary for the ruled surface U to posses tangent planes that are not everywhere isotropic.

Besides the vectors $\mathbf{e}_1(t), \ldots, \mathbf{e}_m(t)$ we consider also their derivatives

$$\dot{\mathbf{e}}_1(t), \ldots, \dot{\mathbf{e}}_m(t).$$

We define the asymptotic bundle as generated subspace

$$A(t) = [\mathbf{e}_1(t), \ldots, \mathbf{e}_m(t), \dot{\mathbf{e}}_1(t), \ldots, \dot{\mathbf{e}}_m(t)], \quad t \in I.$$

Let us assume

$$\dim A(t) = m + l, \quad 0 \le l \le m.$$

By the isotropic Gram-Schmidt orthogonalization process ([9]), we obtain an isotropic orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}\}$ in $A(t)$.

Generally, the following holds

$$\dot{\mathbf{e}}_i = \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j + \sum_{j=1}^{l} \sigma_i^j \mathbf{a}_{m+j}, \quad i = 1, \ldots, m.$$

However, it is possible to construct a more convenient basis for the generating space $E_m(t)$.

THEOREM 1 (THE NATURAL BASIS). *There exists a subinterval $J \subset I$ such that in each generating space $E_m(t)$, $t \in J$, of a regular $(m+1)$-ruled surface*

$U \subset I_n^k$ there exists an orthonormal basis $\{\mathbf{e}_1(t), \dots, \mathbf{e}_m(t)\}$ of $E_m(t)$ which satisfies

$$\tag{1} \begin{aligned} \dot{\mathbf{e}}_i(t) &= \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j(t) + \kappa^i \mathbf{a}_{m+i}(t), \ \ i = 1, \dots, l, \\ \dot{\mathbf{e}}_i(t) &= \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j(t), \ \ i = l+1, \dots, m, \end{aligned}$$

where the functions κ_i (called curvatures) satisfy $\kappa^i \neq 0$, $i = 1, \dots, l$.

The orthonormal basis $\{\mathbf{e}_1(t), \dots, \mathbf{e}_m(t)\}$ of $E_m(t)$ determines uniquely the orthonormal basis $\{\mathbf{e}_1(t), \dots, \mathbf{e}_m(t), \mathbf{a}_{m+1}(t), \dots, \mathbf{a}_{m+l}(t)\}$ of $A(t)$.

PROOF. Let us suppose that in $E_m(t)$, $t \in I$, an orthonormal basis is given which satisfies the conditions mentioned before i.e., such that the vectors $\mathbf{e}_1(t), \dots, \mathbf{e}_\mu(t)$, $\mu = m$ or $\mu \in \{n-k-1, n-k\}$, are Euclidean, and that among the isotropic vectors there are either isotropic vectors of all degrees $\in \{1, \dots, r\}$ or the isotropic degree i is missing.

First, let us construct the required basis of $E_m(t)$ from the Euclidean vectors $\mathbf{e}_1(t), \dots, \mathbf{e}_\mu(t)$. It is constructed as in [2] by means of the isotropic scalar product. In this way we define the vectors $\hat{\mathbf{e}}_1(t), \dots, \hat{\mathbf{e}}_\mu(t)$ which are mutually orthogonal, for $t \in I' \subset I$, and the following holds $\lambda_1 \geq \lambda_2 \geq \cdots \geq \geq \lambda_\mu \geq 0$, where $\lambda_i = ||\hat{\mathbf{e}}_i||^2$, $i = 1, \dots, \mu$.

However, it can happen, that beginning from some $\nu \in \{0, \dots, \mu\}$, the values $\lambda_{\nu+1} = \ldots = \lambda_\mu = 0$. This means that the projections of the vectors $\hat{\mathbf{e}}_{\nu+1}, \dots, \hat{\mathbf{e}}_\mu$ onto U_1 are zero-vectors, i.e., that these vectors are not Euclidean. Now, from these vectors we construct mutually orthogonal isotropic vectors. Let

$$\mathbf{e}(t) = \sum_{i=\nu+1}^{\mu} \gamma^i(t) \mathbf{e}_i(t).$$

Hence, for the vector $\hat{\mathbf{e}}$

$$\tag{2} \hat{\mathbf{e}}(t) = \dot{\mathbf{e}}(t) - \sum_{i=1}^{m} \alpha^i \mathbf{e}_i(t)$$

where

$$\alpha^i = \langle \dot{\mathbf{e}}, \mathbf{e}_i \rangle, \ \ i = 1, \dots, \mu,$$

$$\alpha^{\mu+j} = \dot{e}^{n-k+j} - \sum_{i=1}^{\mu+j-1} \alpha^i e_i^{n-k+j}, \ \ j = 1, \dots, i-1,$$

$$\alpha^{\mu+j} = \dot{e}^{n-k+j+1} - \sum_{i=1}^{\mu+j-1} \alpha^i e_i^{n-k+j+1}, \ \ j = i, \dots, m-\mu,$$

we have

$$\hat{\mathbf{e}}(t) = \sum_{i=\nu+1}^{\mu} \gamma^i(t) \hat{\mathbf{e}}_i(t).$$

By e_i^{n-k+j} we have denoted the $(n-k+j)$-coordinate of the vector $\mathbf{e}_i$ with respect to the basis $\mathbf{b}_{n-k+1}, \dots, \mathbf{b}_n$ of U_2. Functions α^i denote either the components defined by isotropic scalar product which appear when expanding $\dot{\mathbf{e}}$ in basis $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ or are obtained inductively for the coordinates of $\dot{\mathbf{e}}$ in U_2. If in the isotropic part of the basis of $E_m(t)$ the vector of isotropic degree i is not missing, then only the first set of formulas for $\alpha_{\mu+j}$ is applied.

We have already noticed that the vectors $\hat{\mathbf{e}}_{\nu+1}, \dots, \hat{\mathbf{e}}_\mu$ are not Euclidean. Similarly it can be shown that they are neither isotropic vectors of degree $j \in \{1, \dots, r\} \setminus \{i\}$. The index i is excluded only in the case when among the isotropic vectors of the basis of $E_m(t)$ there is an i-isotropic vector.

Let us now determine the extreme values of the ith isotropic coordinate of the vector $\hat{\mathbf{e}}$ (only in the case when there is no i-isotropic vector in the basis of $E_m(t)$) under the condition $\mathbf{e}^2 = \sum_{i=\nu+1}^{\mu} (\gamma^i)^2 = 1$ which guarantees that the extrema exist. We consider the function

$$F(t, \gamma_{\nu+1}, \dots, \gamma_\mu) = \sum_{j=\nu+1}^{\mu} \gamma^j(t) \hat{e}_j^{n-k+i}(t) - \frac{1}{2}\lambda \left(\sum_{j=\nu+1}^{\mu} (\gamma^j)^2(t) - 1 \right).$$

The necessary conditions for the extrema are

$$\frac{\partial F}{\partial \gamma_j} = \hat{e}_j^{n-k+i} - \lambda \gamma^j = 0, \ \ j = \nu+1, \dots, \mu.$$

Hence, the critical points are

$$\gamma^j = \frac{\hat{e}_j^{n-k+i}}{\lambda}, \ \ j = \nu+1, \dots, \mu.$$

In some generating space $E_m(t_0)$, $t_0 \in I'$, we can assume that $\mathbf{e}_{\nu+1}(t_0)$ is the vector of the maximum of the considered function. Therefore

$$\gamma^{\nu+1} = 1, \gamma^{\nu+2} = \ldots = \gamma^{\mu} = 0,$$
$$\lambda_{\nu+1}^{(1)} := \hat{e}_{\nu+1}^{n-k+i} = [\hat{\mathbf{e}}_{\nu+1}]_i,$$
$$\lambda_{\nu+2}^{(1)} := \hat{e}_{\nu+2}^{n-k+i} = 0, \ldots, \lambda_{\mu}^{(1)} := \hat{e}_{\mu}^{n-k+i} = 0,$$

where $[\ ]_i$ denotes the ith range of the given vector. Therefore the vector $\hat{\mathbf{e}}_{\nu+1}$ is i-isotropic, while the vectors $\hat{\mathbf{e}}_{\nu+2}, \ldots, \hat{\mathbf{e}}_{\mu}$ are isotropic of degree $\geq r+1$.

Again, there exists an open interval $I_1 \subset I'$ where the previous equations hold.

If $\lambda_{\nu+1}^{(1)} = 0$, then the ith isotropic coordinate of the vector $\hat{\mathbf{e}}$ is equal to 0. Therefore, in the basis of $A(t)$ there is no i-isotropic vector. We consider the $(r+1)$st isotropic coordinate and repeat the procedure. If there exists some $j_1 \in \{n-k+r+1, \ldots, n\}$ such that $\lambda_{\nu+j_1}^{(1)} := \hat{e}_{\nu+1}^{n-k+j_1} \neq 0$, than j_1-isotropic vector $\hat{\mathbf{e}}_{\nu+1}$ is the vector in the basis of $A(t)$. If all $\lambda_{\nu+1}^{(1)} = 0$, than all the vectors $\hat{\mathbf{e}}_{\nu+1}, \ldots, \hat{\mathbf{e}}_{\mu}$ are zero vectors.

If $\mu = \nu + 2$, then the vector $\mathbf{e}_{\nu+2}$ of the Euclidean part of the basis of $E_m(t)$ is uniquely determined, up to a sign. The corresponding $\hat{\mathbf{e}}_{\nu+2}$ is an isotropic vector of degree $> j_1$.

If $\mu > \nu + 2$, we repeat the previous procedure on vectors $\mathbf{e}_{\nu+2}, \ldots, \mathbf{e}_{\mu}$ and obtain a j_2-isotropic vector $\hat{\mathbf{e}}_{\nu+2}$ of the basis for $A(t)$. Finally, in the same manner, we construct an interval $J \subset \ldots \subset I_1$ and the other vectors $\hat{\mathbf{e}}_{\nu+3}, \ldots, \hat{\mathbf{e}}_{\rho}$ of the basis of $A(t)$, $t \in J$. For the rest of the vectors $\mathbf{e}_{\rho+1}, \ldots, \mathbf{e}_{\mu}$ we get that the vectors $\hat{\mathbf{e}}_{\rho+1}$, …, $\hat{\mathbf{e}}_{\mu}$ are zero-vectors.

It is easy to see that the vectors $\hat{\mathbf{e}}_i$, $\mathbf{e}_j$, which are Euclidean, are mutually orthogonal, which with the previous part implies that $\{\mathbf{e}_1, \ldots, \mathbf{e}_{\mu}, \hat{\mathbf{e}}_1, \ldots, \hat{\mathbf{e}}_{\rho}\}$ is the orthogonal set of vectors.

Finally, let us construct the vectors of the basis of $A(t)$, $t \in J$, from the isotropic vectors $\mathbf{e}_{\mu+1}, \ldots, \mathbf{e}_m$. The vectors $\hat{\mathbf{e}}_{\mu+1}, \ldots, \hat{\mathbf{e}}_m$ defined by (2) are isotropic vectors of degree $\in \{r+1, \ldots, k\} \cup \{i\}$. Together with the vectors obtained from the Euclidean vectors, these vectors span $A(t)$. The same isotropic degree of these vectors can appear either if one of the vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ is Euclidean and the other isotropic, or if both are isotropic (the case when both are Euclidean is treated in the previous part of the proof). Suppose that $\mathbf{e}_E$ is Euclidean vector and $\mathbf{e}_I$ is isotropic of degree i such that

$\hat{\mathbf{e}}_E$ and $\hat{\mathbf{e}}_I$ are isotropic of degree j. Then instead of the vector $\mathbf{e}_E$ in the basis of $E_m(t)$ we take the vector

$$\mathbf{e}_E + \alpha \mathbf{e}_I, \ \ \alpha = -\frac{e_E^{n-k+j}}{e_I^{n-k+j}}$$

which is again unit Euclidean vector, orthogonal to other Euclidean vectors, such that $\widehat{\mathbf{e}_E + \alpha \mathbf{e}_I}$ is of isotropic degree $> j$. Similarly we treat the case of two isotropic vectors $\mathbf{e}_{I1}, \mathbf{e}_{I2}$ of isotropic degrees $I_1 < I_2$. In the basis of $E_m(t)$ instead of the vector $\mathbf{e}_{I1}$ we take the vector $\mathbf{e}_{I1} + \alpha \mathbf{e}_{I2}$ which is again unit isotropic of degree I_1. In that way we orthogonalize the obtained vectors $\hat{\mathbf{e}}_{\mu+1}, \dots, \hat{\mathbf{e}}_m$ in the unique way and obtain the required orthonormal basis for $E_m(t)$, $t \in J$.

Finally, in order to get that the formulas (1) hold, we rename the obtained vectors by permuting their indices. By normalizing the vectors $\hat{\mathbf{e}}_1(t), \dots, \hat{\mathbf{e}}_l(t)$ we define the required unit vectors

$$\mathbf{a}_{m+1}(t), \dots, \mathbf{a}_{m+l}(t)$$

of the basis of $A(t)$. ∎

3. The Striction Space

The tangent $(m+1)$-plane of a regular $(m+1)$-ruled surface U in a point $P(t; v_1, \dots, v_m)$ of the generating space $E_m(t)$ is spanned by

$$[\mathbf{x}_t, \mathbf{x}_1, \dots, \mathbf{x}_m]_P = [\dot{\mathbf{c}} + \sum_{i=1}^{m} v_i \dot{\mathbf{e}}_i, \mathbf{e}_1, \dots, \mathbf{e}_m]_P.$$

The tangent $(m+1)$-planes of U in all points of one fixed generating space $E_m(t)$ lie in

$$T(t) = [\dot{\mathbf{c}}, \dot{\mathbf{e}}_1, \dots, \dot{\mathbf{e}}_m, \mathbf{e}_1, \dots, \mathbf{e}_m].$$

$T(t)$ is called the tangent bundle of U in $E_m(t)$. The asymptotic bundle of U in $E_m(t)$ is a subspace of the tangent bundle $T(t)$ spanned by the tangent $(m+1)$-planes at points at infinity of $E_m(t)$. Therefore

$$m + l \leq \dim T(t) \leq m + l + 1.$$

We distinguish two cases:

(a) $\dim T(t) = m + l$,

(b) $\dim T(t) = m + l + 1$.

CASE (A). In this case there exists a base curve $\mathbf{c}$ such that

$$\dot{\mathbf{c}}(t) \in [\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l}].$$

We can write

$$\dot{\mathbf{c}} = \sum_{i=1}^{m} \tau^i \mathbf{e}_i + \sum_{j=1}^{l} \sigma^j \mathbf{a}_{m+j}.$$

The tangent $(m+1)$-plane in a point $P(v_1, \dots, v_m) \in E_m(t)$ is spanned by the vectors $\mathbf{e}_1, \dots \mathbf{e}_m$ and by

$$\mathbf{x}_t = \dot{\mathbf{c}} + \sum_{i=1}^{m} v^i \dot{\mathbf{e}}_i = \sum_{i=1}^{m} \left(\tau^i + \sum_{j=1}^{m} \alpha_i^j v_j \right) \mathbf{e}_i + \sum_{j=1}^{l} \left(\sigma^j + v^j \kappa^j \right) \mathbf{a}_{m+j}.$$

Hence we can conclude that in the points of $E_m(t)$ which satisfy

$$\sigma^j + v^j \kappa^j = 0, \quad j = 1, \dots, l, \tag{3}$$

the tangent $(m+1)$-plane does not exist. Since $\kappa^i \neq 0$, $i = 1, \dots, l$, the points that satisfy (3) form a $(m-l)$-dimensional subspace $K_{m-l}(t)$ of the generating space $E_m(t)$. This subspace is called the *space of regression*.

Specially, if a base curve is such that $\mathbf{c}(t)$ is a point of a space of regression, then

$$\dot{\mathbf{c}} = \sum_{i=1}^{m} \tau^i \mathbf{e}_i,$$

i.e., $\mathbf{c}$ is tangent to the generating space $E_m(t)$. Therefore the following theorem holds:

THEOREM 2. *If for a generating space $E_m(t)$ the tangent bundle $T(t)$ and the asymptotic bundle $A(t)$ coincide, then there exists a subspace $K_{m-l}(t)$ of $E_m(t)$ in whose points the tangent vector $\dot{\mathbf{c}}(t)$ of the base curve belongs to $E_m(t)$ and the tangent $(m+1)$-planes of the surface does not exist.*

Among the tangent surfaces in I_n^k there appear specially tangent surfaces of a curve lying in some non-isotropic as well as in isotropic j-plane, $j \in \in \{1, \dots, n-1\}$.

CASE (B). In this case there exists a base curve $\mathbf{c}$ such that

$$\dot{\mathbf{c}} \notin [\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l}].$$

Therefore there exists a unit vector $\mathbf{a}_{m+l+1}$ which extends the orthonormal basis of $A(t)$ to an orthonormal basis of $T(t)$ in $E_m(t)$. Hence the vector $\dot{\mathbf{c}}$ can be written in the form

$$\dot{\mathbf{c}} = \sum_{i=1}^{m} \tau^i \mathbf{e}_i + \sum_{j=1}^{l} \sigma^j \mathbf{a}_{m+j} + \sigma^{l+1} \mathbf{a}_{m+l+1}, \quad \sigma^{l+1} \neq 0.$$

The tangent $(m+1)$-plane in a point $P(v_1, \dots, v_m) \in E_m(t)$ is spanned by the vectors $\mathbf{e}_1, \dots, \mathbf{e}_m$ and by

$$\mathbf{x}_t = \dot{\mathbf{c}} + \sum_{i=1}^{m} v^i \dot{\mathbf{e}}_i = \sum_{i=1}^{m} \left(\tau^i + \sum_{j=1}^{m} \alpha_i^j v^j \right) \mathbf{e}_i + \sum_{j=1}^{l} \left(\sigma^j + v^j \kappa^j \right) \mathbf{a}_{m+j} + \sigma^{l+1} \mathbf{a}_{m+l+1}.$$

DEFINITION 2. A point P of a generating space $E_m(t)$ is called the striction (central) point if the tangent $(m+1)$-plane of a surface U in P is j-isotropic, for some $j \in \{1, \dots, k\}$.

CASE I. Let $m + 1 \leq n - k$.

The tangent $(m+1)$-plane is non-isotropic if it is spanned by $m+1$ Euclidean vectors. It is j-isotropic, for some $j \in \{1, \dots, k\}$, if and only if there exists an isotropic vector of degree $\rho \in \{1, \dots, k\}$ parallel to it.

1. Let $\mathbf{a}_{m+l+1}$ be Euclidean.

Since the vectors $\mathbf{e}_1, \dots, \mathbf{e}_m$, $\mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l+1}$ are linearly independent, and $\sigma_{l+1} \neq 0$, the vector $\mathbf{x}_t$ is Euclidean. Hence, the tangent $(m+1)$-plane is non-isotropic in every point of $E_m(t)$. In this case, the striction point does not exist.

In this case if among the vectors $\mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l}$ there are no Euclidean, we can conclude

$$\pi_1(\dot{\mathbf{e}}_i(t)) = \sum_{j=1}^{m} \alpha_i^j \pi_1(\mathbf{e}_j(t)), \quad i = 1, \dots, m.$$

Thus $\pi_1(\mathbf{e}_1), \dots, \pi_1(\mathbf{e}_m)$ span a fixed m-plane in U_1. Therefore, the generators $\mathbf{e}_1, \dots, \mathbf{e}_m$ are parallel to a k-isotropic $(m+k)$-plane in the space I_n^k. Such a ruled surface is called $(m+k)$-conoidal (there exists a $(m+k)$-plane such that the generators are parallel to it) with generators parallel to a k-isotropic plane.

The surfaces of type 3 in I_3^1 are example of such surfaces.

2. Let $\mathbf{a}_{m+l+1}$ be a ρ-isotropic vector, for some $\rho \in \{1, \ldots, k\}$.

If all the vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ are Euclidean, i.e., if the asymptotic bundle is non-isotropic, then the tangent $(m+1)$-plane is j-isotropic, for some $j \in \{1, \ldots, k\}$ if and only if

$$\sigma^j + v^j \kappa^j = 0, \;\; j = 1, \ldots, l. \tag{4}$$

Since $\kappa^j \neq 0, j = 1, \ldots, l$, the previous l equations for the variables $v_1, \ldots, v_m$ determine a $(m-l)$-dimensional subspace $Z_{m-l}(t) \subset E_m(t)$ of striction points. This subspace is called the *striction (central) space.*

Specially, in the case $l = m$ there exists a unique striction point in every generating space $E_m(t)$. The set of all striction points, $t \in J$, is called the *striction (central) curve.*

The surfaces of type 1 and 2 in I_3^1 are examples of such surfaces.

If among the vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ there exist vectors that are not Euclidean (including the possibility that none of these vectors is Euclidean), then the tangent $(m+1)$-plane is j-isotropic, for some $j \in \{1, \ldots, k\}$ in the points of some $(m - l')$-dimensional subspace of $E_m(t)$, $0 \leq l' \leq l$. In particular, in the case $l' = 0$, tangent $(m+1)$-plane is j-isotropic in every point of the surface. But such surfaces do not fit in the notion of an admissible surface ([1], [6], [7]), which has a non-isotropic tangent plane everywhere except along the striction space, and therefore they will be excluded from the further study.

In I_3^1 such surfaces do not exist.

Case II. Let $m + 1 = n - k + r$, $r \in \{1, \ldots, k - 1\}$.

In this case, a $(m + 1)$-plane is non-isotropic if it is spanned by $n - k$ Euclidean vectors, one 1-isotropic, …, one r-isotropic vector. A $(m+1)$-plane is j-isotropic, for some $j \in \{1, \ldots, k\}$, if there exists a ρ-isotropic vector, $\rho \in \{r + 1, \ldots, k\}$ parallel to it. Again, we distinguish the following cases:

1. If $\mathbf{a}_{m+l+1}$ is ρ-isotropic, for some $\rho \in \{r + 1, \ldots, k\}$, and the vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ are Euclidean or $\bar{\rho}$-isotropic, $\bar{\rho} \in \{1, \ldots, r\}$, that is, if $A(t)$ is non-isotropic, then there exists the $(m - l)$-dimensional striction space Z_{m-l} defined by the system (4).

The surfaces of type A,B in I_3^2 are examples of such surfaces.

As before, the surfaces satisfying that among the vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ there are vectors that are $\bar{\rho}$-isotropic, $\bar{\rho} \in \{r + 1, \ldots, k\}$, are excluded from the further study.

Such surfaces do not exist in I_3^2.

2. If $\mathbf{a}_{m+l+1}$ is Euclidean or ρ-isotropic, $\rho \in \{1, \ldots, r\}$ then the striction space in $E_m(t)$ does not exist.

In this case, when the vectors $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ are neither Euclidean nor j-isotropic, $j \in \{1, \ldots, r\}$, then for the projections $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_m$ of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ onto the space $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+r}$ we have

$$\dot{\tilde{\mathbf{e}}}_i = \sum_{j=1}^{m} \alpha_i^j \tilde{\mathbf{e}}_j, \quad i = 1, \ldots, m.$$

Hence $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_m$ span a fixed m-plane in $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+r}$, and the surface is $(m + k - r)$-conoidal with generators parallel to some isotropic plane.

The surfaces of type C and D in I_3^2 are examples of such surfaces.

Let us summarize:

THEOREM 3. *Two main types of skew ruled* $(m + 1)$*-surfaces in the space* I_n^k *are described as follows:*

(a) if $m + 1 \leq n - k$ *and* $\mathbf{a}_{m+l+1}$ *is Euclidean or* $m + 1 = n - k + r$ *and* $\mathbf{a}_{m+l+1}$ *is Euclidean or* ρ*-isotropic,* $\rho \in \{1, \ldots, r\}$*, then the surfaces is a ruled surface without the striction space;*

(b) if $m + 1 \leq n - k$ *and* $\mathbf{a}_{m+l+1}$ *is* ρ*-isotropic,* $\rho \in \{1, \ldots, k\}$ *or* $m + 1 = n - k + r$ *and* $\mathbf{a}_{m+l+1}$ *is* ρ*-isotropic,* $\rho \in \{r + 1, \ldots, k\}$*, while the asymptotic bundle* $A(t)$ *is non-isotropic, then the surface is a ruled surface with the striction space.*

Among surfaces without the striction space, we distinguish the following subtype:

(a1) if $m + 1 \leq n - k$ *and none of* $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ *is Euclidean or* $m + 1 = n - k + r$ *and none of* $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_{m+l}$ *is Euclidean nor* j*-isotropic,* $j \in \{1, \ldots, r\}$*, then the surface is a conoidal surface with generators parallel to an isotropic plane.*

For the $(m + 1)$-ruled surfaces U we define the ith parameter of distribution by

$$\delta_i = \frac{\sigma_{l+1}}{\kappa_i}, \quad i = 1, \ldots, l$$

which is invariant under an admissible transformation of parameters.

The parameter of distribution of 2-ruled surfaces in I_3^1, I_3^2 coincides with the first parameter of distribution here defined.

Let now U be a $(m+1)$-ruled surface such that the isotropic degrees of vectors of the natural basis are constant (on an open interval J) and $\dim T(t) = m + l + 1$. Let the orthonormal basis of tangent bundle $T(t)$ for $t \in J$

$$\{\mathbf{e}_1, \dots \mathbf{e}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l+1}\}$$

be extended to an orthonormal basis of I_n^k

$$\{\mathbf{e}_1, \dots \mathbf{e}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_{m+l+1}, \mathbf{a}_{m+l+2}, \dots, \mathbf{a}_n\}.$$

In such a way we have obtained an associated n-frame of an $(m+1)$-ruled surface U. The following expressions for the derivatives hold

$$\dot{\mathbf{e}}_i(t) = \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j(t) + \kappa^i \mathbf{a}_{m+i}(t), \quad i = 1, \dots, l,$$

$$\dot{\mathbf{e}}_i(t) = \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j(t), \quad i = l+1, \dots, m,$$

$$\dot{\mathbf{a}}_{m+i} = -\epsilon\kappa^i \mathbf{e}_i + \sum_{\mathbf{e}_j = \text{isotropic}} \lambda_i^j \mathbf{e}_j + \sum_{j=1}^{l} \beta_i^j \mathbf{a}_{m+j} + \omega^i \mathbf{a}_{m+l+1} + \sum_{j=2}^{n-m-l} \eta_i^j \mathbf{a}_{m+l+j}, \quad i = 1, \dots, l,$$

$$\dot{\mathbf{a}}_{m+l+1} = \sum_{\mathbf{e}_j = \text{isotropic}} \lambda_{l+1}^j \mathbf{e}_j + \sum_{j=1}^{l} \bar{\omega}^j \mathbf{a}_{m+j} + \sum_{j=1}^{n-m-l} \vartheta^j \mathbf{a}_{m+l+j},$$

$$\dot{\mathbf{a}}_{m+l+i} = \sum_{\mathbf{e}_j = \text{isotropic}} \lambda_{l+i}^j \mathbf{e}_j + \sum_{j=1}^{l} \bar{\eta}_i^j \mathbf{a}_{m+j} + \bar{\vartheta}_i \mathbf{a}_{m+l+1} + \sum_{j=2}^{n-m-l} \vartheta_i^j \mathbf{a}_{m+l+j}, \quad i = 2, \dots, n-m-l,$$

where

$$\epsilon = \begin{cases} 1, & \text{if } \mathbf{a}_{m+i}, \mathbf{e}_i \text{ and their derivatives are Euclidean} \\ 0, & \text{otherwise,} \end{cases}$$

$$\alpha_i^j = -\alpha_j^i,\ \beta_i^j = -\beta_j^i,\ \bar{\eta}_i^j = -\eta_j^i,\ \bar{\omega}^i = -\omega^i,\ \bar{\vartheta}^i = -\vartheta^i,\ \vartheta_i^j = -\vartheta_j^i,$$

if the corresponding vectors and their derivatives are Euclidean.

The striction space $Z_{m-l}(t) \subset E_m(t)$ is given by l equations for the variables $v_1, \dots, v_m$. In particular, in the case $l = m$ there exists a unique striction point in $E_m(t)$. The set of all striction points for $t \in J$ determines the striction curve s. It is given by

$$\mathbf{s}(t) = \mathbf{c}(t) - \sum_{i=1}^{m} \frac{\sigma^i}{\kappa^i}(t)\mathbf{e}_i(t), \quad t \in J.$$

The base curve $\mathbf{c}$ is the striction curve if and only if $v^1 = \dots = v^m = 0$, i.e., if and only if $\sigma^1 = \dots = \sigma^m = 0$. Hence, if the base curve is the striction curve then

$$\dot{\mathbf{s}} = \sum_{i=1}^{m} \tau^i \mathbf{e}_i + \sigma^{m+1}\mathbf{a}_{2m+1}. \tag{5}$$

If the striction curve s is an admissible curve parametrized by the arc length, then the coefficients $\tau^1, \dots, \tau^\rho$ are the direction cosines of the tangent vector $\dot{\mathbf{s}}$

$$\tau^i = \cos\gamma^i, \quad i = 1, \dots, \rho,$$

where $\gamma^i = \angle(\dot{\mathbf{s}}, \mathbf{e}_i)$ is the angle between the projections of the corresponding vectors onto U_1, and $\mathbf{e}_1, \dots, \mathbf{e}_\rho$ are the only (because $m = l$) Euclidean vectors of the natural basis.

The coefficients $\gamma^{\rho+1}, \dots \gamma^m$ represent the isotropic angles

$$\gamma^{\rho+i} = \angle\left(\dot{\mathbf{s}}, \sum_{i=1}^{\rho+i-1} \tau^i \mathbf{e}_i\right), \quad i = 1, \dots, m-\rho.$$

Finally, the following is also true

$$\sigma^{m+1} = \angle\left(\dot{\mathbf{s}}, \sum_{i=1}^{m} \tau^i \mathbf{e}_i\right).$$

The angles γ^i, $i = 1, \dots, m$, σ^{m+1} are called the ith strictions of $(m+1)$-ruled surface U.

In the case when $l < m$, the striction space Z_{m-l} determines a $(m - l + 1)$-ruled surface. This surface is called the *striction (central) surface.* Its parametrized position vector is given by

$$\mathbf{x}(t; v^{l+1}, \dots, v^m) = \mathbf{c}(t) - \sum_{i=1}^{l} \frac{\sigma^i}{\kappa^i}(t)\mathbf{e}_i(t) + \sum_{i=l+1}^{m} v^i \mathbf{e}_i.$$

In the points of the striction space the tangent $(m-l+1)$-plane is spanned by

$$(6)\qquad \mathbf{x}_t = \sum_{i=1}^{m}\left(\bar{\tau}^i - \sum_{j=1}^{l}\alpha_i^j\frac{\sigma^j}{\kappa^j} + \sum_{j=l+1}^{m}\alpha_i^j v^j\right)\mathbf{e}_i + \sigma^{l+1}\mathbf{a}_{m+l+1},$$

$$\mathbf{x}_{l+1} = \mathbf{e}_{l+1}, \dots, \mathbf{x}_m = \mathbf{e}_m,$$

where

$$\bar{\tau}^i = \begin{cases} \tau^i - \left(\frac{\sigma^i}{\kappa^i}\right)^{\cdot}, & \text{if } \ i = 1, \dots, l, \\ \tau^i, & \text{if } \ i = l+1, \dots, m. \end{cases}$$

Since the vector $\mathbf{a}_{m+l+1}$ is ρ-isotropic, $\rho \in \{1, \dots, k\}$, when $m+1 \le n-k$ or $\rho \in \{r+1, \dots, k\}$ when $m+1 = n-k+r$ in the case when the striction space exists, the projection of the tangent $(m-l+1)$-plane onto U_1 or onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+r}$ is contained in m-plane spanned by the projections of the vector $\mathbf{e}_1, \dots, \mathbf{e}_m$. Therefore, if $l = 1$, the projection of the striction surface is the envelope of the projection of the family of generators. These surfaces are generalization of the surfaces of type 1 in I_3^1 and of type A in I_3^2. Therefore, the following theorem holds:

THEOREM 4. *The projection of the striction m-surface onto U_1 ($U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+r}$) is the envelope of the projections of the family of the generators.*

4. 2-Ruled Surfaces in I_n^k

2-ruled surfaces are given by

$$\mathbf{x}(t, v) = \mathbf{c}(t) + v\mathbf{e}(t).$$

Let us consider the case of non-cylindrical skew surfaces, i.e., $m = l = 1$. Theorems 2, 4 and the previous considerations imply that there are exactly the following types of 2-ruled surface in I_n^k:

COROLLARY 1. *Among the surfaces without the striction curve there are exactly two types:*

(a1) conoidal surfaces with generators parallel to some isotropic plane;

(a2) if $\dim U_1 \ge 3$, surfaces having both tangent and asymptotic plane non-isotropic.

Among the surfaces with the striction curve there are exactly two types:

(b1) surfaces whose projections of the striction curve to U_1 *(*$U_1 \oplus \mathbf{b}_1$*) degenerates to a point;*

(b2) surfaces whose projections of the striction curve is the envelope of the projection of the generators.

More precisely we have:

I. $m + 1 = 2 \leq n - k$, i.e., $\dim U_1 \geq 2$.

In this case, the generator is determined by a unit Euclidean vector $\mathbf{e}$. The vector $\mathbf{a}$ is determined by

$$\dot{\mathbf{e}} = \kappa \mathbf{a},$$

where $\kappa = ||\dot{\mathbf{e}}||$, if $\dot{\mathbf{e}}$ is Euclidean, or $\kappa = [\dot{\mathbf{e}}]_j$, if $\dot{\mathbf{e}}$ is j-isotropic, for some $j \in \{1, \ldots, k\}$. $[\ ]_j$ denotes the jth range.

There exists the striction point in $E_1(t)$ if and only if $\mathbf{a}_3$ is some ρ-isotropic vector, $\rho \in \{1, \ldots, k\}$. This point is determined by $\sigma^1 + \nu\kappa^1 = 0$. The set of all striction points determines the striction curve.

A base curve c is the striction curve if and only if $\nu = 0$. If c parametrized by the arc length, then

$$\dot{\mathbf{c}} = \mathbf{e} + \sigma^2 \mathbf{a}_3,$$

or, in the projection onto U_1

$$\pi_1(\dot{\mathbf{c}}) = \pi_1(\mathbf{e}).$$

The projection of the striction curve is therefore the envelope of the projection of the family of generators. Such surfaces are the surfaces of type 1 in I_3^1 and of type A in I_3^2.

The striction curve c can also be degenerated, i.e., $\pi_1(\mathbf{c})$ can be a point. Such surfaces are the surfaces of type 2 in I_3^1 and of type B in I_3^2.

The parameter of distribution of 2-ruled surface with the striction curve is defined as

$$\delta = \frac{\sigma^2}{\kappa},$$

where

$$\sigma^2 = \angle(\dot{\mathbf{c}}, \mathbf{e}).$$

The striction curve does not exist if and only if $\mathbf{a}_3$ is an Euclidean vector. If the vector $\mathbf{a}$ is j-isotropic, for some $j \in \{1, \ldots, k\}$, then the projection of $\dot{\mathbf{e}}$ onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+j-1}$ is zero-vector. Therefore, the projection of $\mathbf{e}$ is constant. The generators $\mathbf{e}$ are parallel to some $(k-1)$-isotropic $(k-j+2)$-plane, for $j > 1$, and to k-isotropic $(k+1)$-plane for $j = 1$. Such a surface is $(k-j+2)$-conoidal.

Furthermore, if $\mathbf{e}, \mathbf{a}, \mathbf{a}_3$ are all Euclidean ($\dim U_1 \geq 3$), we get a surface of type $(a2)$.

II. $m+1 = 2 = n-k+r$. This is possible only in the case when $n-k = 1$, $r = 1$, i.e., in the space I_n^{n-1}. The vector $\mathbf{e}$ can be either a unit Euclidean or a unit 1-isotropic vector. In both cases, $\mathbf{a}$ is defined by

$$\dot{\mathbf{e}} = \kappa \mathbf{a}.$$

Furthermore, the striction point exists if and only if $\mathbf{a}_3$ is ρ-isotropic, for some $\rho \in \{2, \dots, k\}$. For such a striction curve we have the same conclusions as in the previous case.

The striction curve does not exist if and only if $\mathbf{a}_3$ is 1-isotropic, in the case when $\mathbf{e}$ is Euclidean, or if and only if $\mathbf{a}_3$ is Euclidean, in the case when $\mathbf{e}$ is 1-isotropic. If the vector $\mathbf{a}$ is j-isotropic, for some $j \in \{2, \dots, k\}$, then the projection of $\dot{\mathbf{e}}$ onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-k+j-1}$ is zero-vector. In these cases the surfaces are $(k-j+2)$-conoidal. More precisely, if $j = 2$ and $\mathbf{e}$ is Euclidean, this k-plane is $(k-1)$-isotropic; if $\mathbf{e}$ is 1-isotropic then it is k-isotropic. If $j \in \{3, \dots, k\}$, this $(k-j+2)$-plane is $(k-1)$-isotropic.

5. $(n-1)$-Ruled Surfaces in I_n^k

We consider the surfaces such that the generating spaces $E_{n-2}(t)$ are $(n-2)$-dimensional and the surfaces are non-cylindrical skew surfaces, i.e., $m = n-2$, $l = 1$. Then the asymptotic bundle is generated by

$$A(t) = [\mathbf{e}_1, \dots, \mathbf{e}_{n-2}, \mathbf{a}_{n-1}]$$

and the following equations for the derivatives hold

$$\dot{\mathbf{e}}_1(t) = \sum_{j=1}^{m} \alpha_1^j \mathbf{e}_j(t) + \kappa^1 \mathbf{a}_{n-1}(t),$$

$$\dot{\mathbf{e}}_i(t) = \sum_{j=1}^{m} \alpha_i^j \mathbf{e}_j(t), \quad i = 2, \dots, n-2.$$

From Theorem 2 it follows immediately:

COROLLARY 2. *There exist exactly the following types of ruled hypersurfaces in* I_n^k*:*

(a) conoidal surfaces with generators parallel to an isotropic plane;

(b) surfaces with the striction space.

More precisely:

I. Let $n - 1 \leq n - k$. This is possible if and only if $k = 1$, i.e., in the space I_n^1. In this case the generators $\mathbf{e}_1, \ldots, \mathbf{e}_{n-2}$ are all Euclidean.

There exists the striction space $Z_{n-3}(t)$ in the generating space $E_{n-2}(t)$ if and only if $\mathbf{a}_n$ is equal to the completely isotropic vector $\mathbf{b}_1$. Then, in the projection onto U_1, the $(n-2)$-striction surface is the envelope of the family of generators.

Obviously, the striction curve exists if and only if $n = 3$.

The striction space does not exist if and only if $\mathbf{a}_n$ is Euclidean. Since the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_{n-2}$ are also Euclidean, the vector $\mathbf{a}_{n-1}$ must be completely isotropic. The projection onto U_1 of the generators span a constant $(n - 2)$-plane. Therefore, generators are parallel to an isotropic hyperplane in I_n^1. Such a surface is $(n - 1)$-conoidal.

II. Let $n-1 = n-k+r$, for some $r \in \{1, \ldots, k-1\}$, i.e., we consider the space I_n^{r+1}. In this case the generators can be either Euclidean or j-isotropic, $j \in \{1, \ldots, r\}$.

Let the generators be such that there are $n-k$ Euclidean and one of each of j-isotropic vectors, $j \in \{1, \ldots, r\} \setminus \{i\}$. The striction space Z_{n-3} exists if and only if $\mathbf{a}_n$ is $(r+1)$-isotropic vector. In projection onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-1}$ the striction surface is the envelope of the family of generators.

The striction space does not exist if and only if $\mathbf{a}_n$ is i-isotropic. Then $\mathbf{a}_{n-1}$ must be $(r + 1)$-isotropic. Now the projection of generators onto the space $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-1}$ is a constant $(n - 2)$-plane. Hence the generators are parallel to a $(k - i)$-isotropic hyperplane. Such a surface is $(n - 1)$-conoidal.

Let the generators be such that there are $n - k - 1$ Euclidean and one of each of j-isotropic vectors, $j \in \{1, \ldots, r\}$. Then, the striction space Z_{n-3} exists if and only if $\mathbf{a}_n$ is again $(r+1)$-isotropic. It does not exist if and only if $\mathbf{a}_n$ is Euclidean. In the case when the striction space exists, in the projection onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-1}$, the striction surface is the envelope of the family of generators.

In the case when it does not exist, $\mathbf{a}_{n-1}$ must be $(r+1)$-isotropic. By projecting onto $U_1 \oplus \mathbf{b}_{n-k+1} \oplus \cdots \oplus \mathbf{b}_{n-1}$ we see that the generators span a constant $(n-2)$-plane. Such a surface is $(n-1)$-conoidal. The hyperplane to which all the generators are parallel is k-isotropic.

References

[1] H. BRAUNER, Geometrie des zweifach isotropen Raumes II., *Jour. f. d. reine u. angew. Math.* **226** (1967), 132–158.

[2] H. FRANK and O. GIERING, Verallgemeinerte Regelflächen, *Math. Z.* **150** (1976), 261–271.

[3] H. FRANK and O. GIERING, Regelflächen mit Zentralräumen, *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II* **187(4–7)** (1978), 139–163.

[4] Ž. MILIN ŠIPUŠ, *Višedimenzionalni izotropni prostori*, Ph. D. Thesis, University of Zagreb, 1999.

[5] Ž. MILIN ŠIPUŠ and B. DIVJAK, Curves in n-dimensional k-isotropic space, *Glasnik Matematički* **Vol.33(53)** (1998), 267–286.

[6] H. SACHS, *Isotrope Geometrie des Raumes*, Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden, 1990.

[7] H. SACHS, *Zur Liniengeometrie isotroper Räume*, Habilitationsschrift a. d. Universität Stuttgart, Privatdruck, 1972.

[8] W. O. VOGEL, Regelflächen im isotropen Raum, *Jour. f. d. reine u. angew. Math.* **202** (1957), 196–214.

[9] H. VOGLER and H. WRESNIK, Endlichdimensionale isotrope Räume I_n^k vom Isotropiegrad k, *Grazer Math. Berichte* Ber. Nr. **307** (1989), 1–46.

Željka Milin Šipuš

Department of Mathematics
University of Zagreb
P.P. 335
10 002 Zagreb
Croatia
milin@math.hr

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 207–210

GRAHAM'S EXAMPLE IS THE ONLY TIGHT ONE FOR $P \,||\, C_{\max}$

By

GYÖRGY DÓSA

(Received December 20, 2004)

1. Introduction

One of the most known and well-studied problem of scheduling theory is $P \,||\, C_{\max}$. There is a fixed set of jobs $\mathcal{J} = \{T_1, \dots, T_n\}$. Each job must be processed exactly one of m identical parallel machines. T_i $(i = 1, \dots, n)$ denotes both the job itself and its processing time which is the same on all machines. If the processing of a job started on a machine then it must be finished without any interruption. One machine can process only one job at the same time. Each job and machine are available at time zero. The goal is to minimize the makespan, i.e., the time when the last job is finished.

A schedule of the set of jobs $\mathcal{J}$ is a partition $\mathcal{P} = \{P_1, P_2, \dots, P_m\}$ which means that part P_i is loaded to machine i. Denote by $\mathcal{L}(\mathcal{P})$ the makespan if partition $\mathcal{P}$ is applied and let $\mathcal{L}(P_i)$ be the load of machine i in this case. Then $\mathcal{L}(\mathcal{P}) = \max\{\mathcal{L}(P_i) \mid i = 1, \dots, m\} = \max\{\sum_{T_k \in P_i} T_k \mid i = 1, \dots, m\}$. A schedule $\mathcal{P}^*$ is optimal if for any other partition $\mathcal{P}$ the inequality $\mathcal{L}(\mathcal{P}^*) \leq \mathcal{L}(\mathcal{P})$ holds. It is obvious that optimal schedule exists as the number of partitions is finite. Throughout the paper the makespan of the optimal schedule is denoted by $C^*(\mathcal{J})$ or simply C^*.

The $P \,||\, C_{\max}$ problem is NP-complete. Therefore several heuristic procedures have been suggested. One of the earliest algorithms is Graham's LPT list scheduling [1]. Let $\mathcal{A}$ be a heuristic algorithm. Denote by $C_{\mathcal{A}}(\mathcal{J})$ the

---

Mathematics Subject Classification (2000): 90B35 makespan of the partition of set $\mathcal{J}$ determined by $\mathcal{A}$ for a given number of machines. Let

$$R_m(\mathcal{A}) = \sup_{\mathcal{J}} \left\{ \frac{C_{\mathcal{A}}(\mathcal{J})}{C^*(\mathcal{J})} \right\}$$

be *the performance ratio* of the algorithm. In [1] Graham introduced the LPT list scheduling algorithm: First we arrange the jobs into a nonincreasing order of their processing times, then the tasks are scheduled in this order individually in such a way that each job is assigned to a machine which has minimal current load.

THEOREM 1 ([1]). $R_m(\text{LPT}) = \frac{4m-1}{3m}$. ∎

Graham's example to show the tightness of the theorem is

$\mathcal{J}^* = \{2m-1, 2m-1, 2m-2, 2m-2, \ldots, m+1, m+1, m, m, m\}$,

i.e., it has two copies of jobs with processing time $l = m+1, m+2, \ldots, 2m-1$, and three copies of jobs with processing time m. The number of jobs is $2m+1$. In the optimal solution the last three jobs are on the same machine. The load of this machine is $3m$. Pairs are formed from the remaining $2m-2$ jobs such that the sum of the two processing time is again $3m$ and these pairs are loaded to the remaining $m-1$ machines. The LPT list scheduling assigns the first $2m$ jobs to the machines that each machine has two jobs and total processing time $3m-1$. The makespan is obtained after the assignment of the last job with length m and is $4m-1$. This note is devoted to show that for any other task-set $\mathcal{J}$ holds that $\frac{C_{\text{LPT}}(\mathcal{J})}{C^*(\mathcal{J})} < \frac{4m-1}{3m}$.

2. There is no other tight example

THEOREM 2. *Let $\mathcal{J} \neq \mathcal{J}^*$. Then $\frac{C_{LPT}(\mathcal{J})}{C^*(\mathcal{J})} < \frac{4m-1}{3m}$.*

REMARK 1. Let us introduce the following notation:

$$R_{m,n}(\mathcal{A}) = \sup_{\mathcal{J}} \left\{ \frac{C_{\mathcal{A}}(\mathcal{J})}{C^*(\mathcal{J})}, |\mathcal{J}| = n \right\}.$$

It is obvious that $R_m(\mathcal{A}) = \sup_n \{R_{m,n}(\mathcal{A})\}$. It is not too hard to prove that $R_{m,n}(\text{LPT})$ is strictly less than $\frac{4m-1}{3m}$, if $n \neq 2m+1$, i.e. $n \neq |\mathcal{J}^*|$.

The tightness of the bound of the performance ratios are shown by the notion of counterexample.

DEFINITION 1. Let $\mathscr{A}$ be a heuristic method. Assume that the number of machines is fixed. Let p, and q be positive numbers with $p > q$. A (p/q) counterexample is a set $\mathscr{J}$ of jobs such that $C_{\mathscr{A}}(\mathscr{J}) = p$, and $C^*(\mathscr{J}) = q$. $\mathscr{J}$ is minimal (p/q) counterexample if it is a counterexample and has minimal number of jobs.

The next definition will be useful.

DEFINITION 2. The moment during the execution of the LPT list scheduling algorithm just before assigning the last task is denoted by *.

LEMMA 1. *Assume that $m \geq 2$, let $\mathscr{J}$ be minimal (p/q) counterexample for the algorithm LPT. Then the following inequality holds: $T_n \geq \frac{m}{m-1}(p-q)$, where T_n is the last job with minimum processing time.*

PROOF. Task T_n was assigned to the earliest possible time, say s. It can be supposed by minimality of $\mathscr{J}$ that before assigning T_n the makespan is less than p, and by assigning T_n it is equal to p. There is a machine among the other ones with completion time not higher than $q - \frac{p-q}{m-1}$. Task T_n has been assigned to the least loaded machine, thus $s \leq q - \frac{p-q}{m-1}$. After assigning the last task the completion time p is reached, thus the processing time of job T_n is at least $p - \left(q - \frac{p-q}{m-1}\right) = \frac{m}{m-1}(p-q)$. ∎

DEFINITION 3. Let $\mathscr{J}$ be a job-set, for which $|\mathscr{J}| \leq 2m$ holds. The schedule of $\mathscr{J}$ is regular, if

1. At most two jobs are scheduled for every machines,
2. If two jobs A_i and B_i are assigned to machine i and $A_i > B_i$, then job A_i precedes job B_i.
3. If two jobs are assigned to machines i, and j, say A_i, B_i and A_j, B_j, resp., and $i < j$, then $A_i \geq A_j$, and $B_i \leq B_j$.

LEMMA 2. *Suppose that there is an optimal schedule such that at most two jobs are scheduled for any machine. Then the regular schedule is optimal and LPT produces the regular schedule.*

PROOF. The statement follows immediately. ∎

PROOF OF THEOREM 2. Suppose that there is a minimal counterexample $\mathscr{J}$ for which $q = 3m$ and $p > 4m - 1$ holds. Then $T_n > m$ according to Lemma 1. Thus, each machine has at most two jobs in the optimal schedule.

This is a contradiction according to Lemma 2, because in such a case the heuristic solution is optimal. Thus $R_m(\text{LPT}) \le \frac{4m-1}{3m}$.

Now suppose that $q = 3m$ and $p = 4m - 1$. Then it follows from Lemma 1 that $T_n \ge m$. If $T_n > m$, then a contradiction is obtained similarly, thus $T_n = m$. Hence it is easy to see that the completion time of the machine to which T_n is assigned by LPT is exactly $3m - 1$ at * therefore the completion times of all other machines are also exactly $3m - 1$. Then the completion time of every optimal machine is exactly $3m$. It follows from Lemma 2 that there is at least one machine of the optimal solution having more than two jobs. From $C^* = 3m$, and $T_n = m$ it follows that the number of these jobs is 3, and the processing time of each of them is m. At moment * there are two machines having a job with processing time m. The processing times of the other job on the same machines are exactly $2m - 1$. Each of these jobs are assigned in the optimal schedule together with a job of processing time $m+1$. These jobs are scheduled at time * together with jobs having processing time $2m - 2$, etc. By induction we get that the counterexample $\mathcal{I}$ must be equal to $\mathcal{I}^*$. ∎

References

[1] R. L. GRAHAM, Bounds on multiprocessor timing anomalies, *SIAM J. Appl. Math.* **17** (1969), 416–429.

György Dósa
Department of Mathematics
University of Veszprém
Veszprém, Hungary
dosagy@almos.vein.hu

ANNALES UNIV. SCI. BUDAPEST., **47** (2004), 211–229

ON QUASILINEAR PARABOLIC FUNCTIONAL DIFFERENTIAL EQUATIONS WITH DISCONTINUOUS TERMS

By

LÁSZLÓ SIMON

(Received February 2, 2005)

Introduction

This work was motivated by works [4]–[6], [9], [10] where nonlinear parabolic functional differential equations of certain types were considered. In [4], [5] M. Chipot, L. Molinet, B. Lovat considered the equation

$$(0.1)\qquad D_t u - \sum_{i,j=1}^{n} D_i[a_{ij}(l(u(\cdot,t))D_j u] + a_0(l(u(\cdot,t))u = f \text{ in } \Omega \times R^+$$

where $\Omega \subset R^n$ is a bounded domain with sufficiently smooth boundary,

$$\sum_{i,j=1}^{n} a_{ij}(\zeta)\xi_i\xi_j \geq \lambda|\xi|^2 \text{ for all } \xi \in R^n, \quad \zeta \in R$$

with some constant $\lambda > 0$,

$$l(u(\cdot,t)) = \int_\Omega g(x)u(x,t)dx$$

with a given function $g \in L^2(\Omega)$. Existence and asymptotic properties (as $t \to \infty$) of solutions of initial-boundary value problems for (0.1) were proved. Such problems arise in diffusion process (for heat or population), where the diffusion coefficient depends on a nonlocal quantity.

In [9], [10] systems of nonlinear parabolic functional problems were considered when modelling diffusion, convection, absorption reaction of chemicals in porous media and reactive transport through an array of cells with semi-permeable membranes.

Finally, in [6] a climate model was considered where the differential equation contained discontinuous and delay terms. Some results on the stabilization (as $t \to \infty$) of solutions to certain nonlinear parabolic functional differential equations with discontinuous and delay terms were proved in [13] by using methods of [7], [8] proving stabilization results for the above mentioned climate model without delay terms.

The aim of the present paper is to consider equations of the form

(0.2)
$$D_t u - \sum_{i=1}^{n} D_i[a_i(t,x,u(t,x),Du(t,x);u)] + a_0(t,x,u(t,x),Du(t,x);u) = f$$

in $Q_T = (0,T)\times\Omega$ with certain homogeneous boundary and initial conditions, where the functions

$$a_i : Q_T \times R^{n+1} \times L^p(0,T;V) \to R$$

(with a closed linear subspace V of the Sobolev space $W^{1,p}(\Omega)$, $2 \le p < \infty$) have certain special forms such that they contain terms which do not depend continuously on u. These equations are generalizations of the parabolic functional differential equations of [4]–[6], [13] and in certain special cases can be considered as models for nonlinear diffusion processes where the diffusion coefficient depends on a nonlocal quantity. The conditions I–V of the existence theorem on (0.2) are generalizations of the above conditions with respect to (0.1).

In Section 1 we shall prove a general existence theorem for (0.2) in the continuous case. In Section 2 a special form of (0.2) with discontinuous terms will be studied. First we prove existence of weak solutions, then we prove the uniqueness of the solution if certain additional conditions are satisfied. Some results on boundedness and stabilization of the solution as $t \to \infty$ will be shown in a separate paper.

1. Existence of solutions

Let $\Omega \subset R^n$ be a bounded domain having the uniform C^1 regularity property (see [1]) and $p \ge 2$ be a real number. Denote by $W^{1,p}(\Omega)$ the usual Sobolev space of real valued functions with the norm

$$\|u\| = \left[\int_\Omega (|Du|^p + |u|^p)\right]^{1/p}$$

and let $V \subset W^{1,p}(\Omega)$ be a closed linear subspace. Denote by $X = L^p(0,T;V)$ the Banach space of the set of measurable functions $u:(0,T)\to V$ such that $\|u\|^p$ is integrable and define the norm by

$$\|u\|^p_{L^p(0,T;V)} = \int_0^T \|u(t)\|^p_V dt.$$

The dual space of $X = L^p(0,T;V)$ is $X^\star = L^q(0,T;V^\star)$ where $1/p + 1/q = 1$ and $V^\star$ is the dual space of V (see, e.g., [11], [14]).

Assume that

I. The functions $a_i : Q_T \times R^{n+1} \times L^p(0,T;V) \to R$ satisfy the Carathéodory conditions for arbitrary fixed $v \in L^p(0,T;V)$ $(i = 0,1,\dots,n)$.

II. There exist bounded (nonlinear) operators $g_1 : L^p(0,T;V) \to R^+$ and $k_1 : L^p(0,T;V) \to L^q(Q_T)$ such that

$$|a_i(t,x,\zeta_0,\zeta;v)| \le g_1(v)\left[|\zeta_0|^{p-1} + |\zeta|^{p-1}\right] + [k_1(v)](t,x), \quad i = 0,1,\dots,n$$

for a.e. $(t,x) \in Q_T$, each $(\zeta_0,\zeta) \in R^{n+1}$ and $v \in L^p(0,T;V)$ and there exists $\delta > 0$ such that k_1 is continuous as a map from $L^p(0,T;W^{1-\delta,p}(\Omega))$ into $L^q(Q_T)$.

III. $\sum_{i=1}^n [a_i(t,x,\zeta_0,\zeta;v) - a_i(t,x,\zeta_0,\zeta^\star;v)](\zeta_i - \zeta_i^\star) > 0$ if $\zeta \ne \zeta^\star$.

IV. There exist a constant $c_2 > 0$ and a bounded operator $k_2 : L^p(0,T;V) \to L^1(Q_T)$ such that

$$\sum_{i=0}^n a_i(t,x,\zeta_0,\zeta;v)\zeta_i \ge c_2\left[|\zeta_0|^p + |\zeta|^p\right] - [k_2(v)](t,x) \quad (k_2 \ge 0) \tag{1.3}$$

for a.e. $(t,x) \in Q_T$, all $(\zeta_0,\zeta) \in R^{n+1}$, $v \in L^p(0,T;V)$,

$$\lim_{\|v\|_X \to \infty} \frac{\|k_2(v)\|_{L^1(Q_T)}}{\|v\|^p_X} = 0 \tag{1.4}$$

and k_2 is continuous as a map from $L^p(0,T;W^{1-\delta,p}(\Omega))$ into $L^1(Q_T)$.

V. If $(u_k) \to u$ strongly in $L^p(0,T;W^{1-\delta,p}(\Omega))$ and $(\zeta_{0,k}) \to \zeta_0$, $(\zeta_k) \to$ $\to \zeta$ then

$$a_i(t,x,\zeta_{0,k},\zeta_k;u_k) \to a_i(t,x,\zeta_0,\zeta;u) \text{ as } k \to \infty \tag{1.5}$$

for a.e. $(t,x) \in Q_T$ $(i = 0,1,\dots,n)$.

Then we may define operator $A : L^p(0, T; V) \to L^q(0, T; V^\star)$ by

$$[A(u), v] =$$

$$\int_{Q_T} \left[\sum_{i=1}^{n} a_i(t, x, u(t, x), Du(t, x); u) D_i v + a_0(t, x, u(t, x), Du(t, x); u) v \right] dt dx,$$

$$u, v \in X = L^p(0, T; V).$$

THEOREM 1.1. *Assume I–V. Then $A : X \to X^\star$ is bounded, demicontinuous, pseudomonotone with respect to $D(L) = \{u \in X : D_t u \in \in X^\star, u(0) = 0\}$ and it is coercive. Consequently (see, e.g.,* [2]*), for any $f \in X^\star$ there exists a (weak) solution $u \in D(L)$ of*

$$D_t u + A(u) = f, \quad u(0) = 0. \tag{1.6}$$

PROOF. a) Boundedness of A (i.e. A maps bounded sets of X into bounded sets of $X^\star$) follows from I, II and Hölder's inequality.

b) We show that A is demicontinuous, i.e. if $(u_k) \to u$ strongly in X then $(A(u_k)) \to A(u)$ weakly in $X^\star$. Assuming that $(u_k) \to u$ strongly in X, we obtain by V for a subsequence

(1.7)

$$a_i(t, x, u_k(t, x), Du_k(t, x); u_k) \to a_i(t, x, u(t, x), Du(t, x); u) \text{ a.e. in } Q_T$$

since for a subsequence $(u_k) \to u$ and $(Du_k) \to Du$ a.e. in Q_T. So, by using II and Hölder's inequality, we can apply Vitali's theorem to obtain that $(A(u_k)) \to A(u)$ weakly in $X^\star$.

c) By using arguments of [3], we show that A is pseudomonotone with respect to $D(L)$, i.e.

$$u_k \in D(L), \quad (u_k) \to u \text{ weakly in } X, \quad (D_t u_k) \to D_t u \text{ weakly in } X^\star, \tag{1.8}$$

$$\limsup_{k \to \infty} [A(u_k), u_k - u] \le 0 \tag{1.9}$$

imply

$$\lim_{k \to \infty} [A(u_k), u_k - u] = 0 \text{ and } (A(u_k)) \to A(u) \text{ weakly in } X^\star. \tag{1.10}$$

Since $W^{1,p}(\Omega)$ is compactly imbedded in $W^{1-\delta,p}(\Omega)$, one may apply a well known imbedding theorem, see [11]. Consequently, (1.8) implies that for a subsequence (denoted for simplicity again by (u_k))

$$(u_k) \to u \text{ strongly in } L^p(0, T; W^{1-\delta,p}(\Omega)) \tag{1.11}$$

and

(1.12) $(u_k) \to u$ in the norm of $L^p(Q_T)$ and a.e. in Q_T.

Define the function p_k by

$$p_k(t,x) = \sum_{i=1}^{n} \left[a_i\,(t,x,u_k(t,x), Du_k(t,x); u_k) - \right. \tag{1.13}$$

$$\left. -a_i\,(t,x,u(t,x), Du(t,x); u)\right] (D_i u_k - D_i u) +$$

$$+ \left[a_0\,(t,x,u_k(t,x), Du_k(t,x); u_k) - a_0\,(t,x,u(t,x), Du(t,x); u)\right] (u_k - u).$$

Assumption IV implies

$$\sum_{i=1}^{n} a_i(t,x,u_k, Du_k; u_k) D_i u_k + a_0(t,x,u_k, Du_k; u_k) u_k \geq \tag{1.14}$$

$$\geq c_2 \left[|u_k|^p + |Du_k|^p\right] - [k_2(u_k)](t,x) \geq c_2 |Du_k|^p - [k_2(u_k)](t,x).$$

By using Young's inequality, it is not difficult to obtain from (1.14) and II

$$p_k(t,x) \geq \tag{1.15}$$

$$\geq \frac{c_2}{2} |Du_k|^p - c_3 \left[|u_k|^p + |Du|^p + |u|^p + |k_1(u_k)|^q + |k_1(u)|^q + k_2(u_k)\right]$$

with some positive constant c_3. The expression in (1.15), which is multiplied by c_3, is convergent in $L^1(Q_T)$ because of (1.11), II, IV, thus it is a.e. convergent and so bounded a.e., for a subsequence. Hence, for a fixed point $(t,x) \in Q_T$

(1.16) $p_k(t,x) < 0$ implies that $(Du_k(t,x))$ is bounded

except possibly a set of zero measure.

One can write $p_k(t,x)$ in the form

$$p_k(t,x) = q_k(t,x) + r_k(t,x) + s_k(t,x) \tag{1.17}$$

where

$$q_k = \sum_{i=1}^{n} \left[a_i(t,x,u_k(t,x), Du_k(t,x); u_k) - \right.$$

$$\left. - a_i(t,x,u_k(t,x), Du(t,x); u_k)\right] (D_i u_k - D_i u),$$

$$r_k = \sum_{i=1}^{n} \left[a_i(t,x,u_k(t,x), Du(t,x); u_k) - \right.$$

$$\left. - a_i(t,x,u(t,x), Du(t,x); u)\right] (D_i u_k - D_i u),$$

$$s_k = \left[a_0(t,x,u_k(t,x), Du_k(t,x); u_k) - a_0(t,x,u(t,x), Du(t,x); u)\right] (u_k - u).$$

Denoting by $p_k^-(t,x)$ the negative part of $p_k(t,x)$ and by χ_k the characteristic function of p_k^-, we find by (1.17)

$$-p_k^- = \chi_k q_k + \chi_k r_k + \chi_k s_k. \tag{1.18}$$

(1.11), (1.12), V and II, (1.16) imply

$$(\chi_k r_k) \to 0 \text{ and } (\chi_k s_k) \to 0 \text{ a.e. in } Q_T. \tag{1.19}$$

Since $\chi_k q_k \geq 0$, $p_k^- \geq 0$, (1.18), (1.19) imply that

$$p_k^-(t,x) \to 0 \text{ a.e. in } Q_T. \tag{1.20}$$

Further, by (1.15)

$$p_k(t,x) \geq -c_3[|u_k|^p + |Du|^p + |u|^p + |k_1(u_k)|^q + |k_1(u)|^q + k_2(u_k)]$$

where the right hand side is convergent in $L^1(Q_T)$ and so it is equiintegrable, which implies that (p_k^-) is equiintegrable. Combining this fact with (1.20), we obtain

$$\lim_{k\to\infty} \int_{Q_T} p_k^- = 0. \tag{1.21}$$

By (1.8), (1.9)

$$\limsup_{k\to\infty} \int_{Q_T} p_k = \limsup_{k\to\infty} [A(u_k), u_k - u] - \lim_{k\to\infty} [A(u), u_k - u] \leq 0, \tag{1.22}$$

thus

$$\limsup_{k\to\infty} \int_{Q_T} p_k^+ = \limsup_{k\to\infty} \int_{Q_T} p_k + \lim_{k\to\infty} \int_{Q_T} p_k^- \leq 0,$$

consequently,

$$(p_k^+) \to 0 \text{ in } L^1(Q_T)$$

and so

$$(p_k) \to 0 \text{ in } L^1(Q_T) \text{ and a.e. in } Q_T \tag{1.23}$$

for a subsequence. Combining (1.23) with (1.15), we find that (Du_k) is bounded a.e. Consequently, $(r_k) \to 0$, $(s_k) \to 0$ a.e. in Q_T which implies that

$$(q_k) \to 0 \text{ a.e. in } Q_T. \tag{1.24}$$

Since (Du_k) is bounded a.e. in Q_T, for a.e. fixed $(t,x) \in Q_T$ is has a convergent subsequence: $(Du_k)(t,x) \to v(t,x)$. Consequently, by (1.11), (1.12) we obtain from V

$$\sum_{i=1}^{n} [a_i(t,x,u(t,x),v(t,x);u) - a_i(t,x,u(t,x),Du(t,x);u)][v_i(t,x) - D_i u(t,x)]=0$$

which implies $v(t,x) = Du(t,x)$ (for a.e. (t,x)) by III, i.e.

$$(Du_k) \to Du \text{ a.e. in } Q_T. \tag{1.25}$$

By (1.22), (1.23)

$$\lim_{k\to\infty} [A(u_k), u_k - u] = 0, \tag{1.26}$$

further, by (1.11), (1.12), (1.25), V, Hölder's inequality, Vitali's theorem we obtain that

$$(A(u_k)) \to A(u) \text{ weakly in } X^\star,$$

i.e. with (1.26) we have proved (1.10).

d) Finally, by IV

$$\frac{[A(v),v]}{\|v\|_X} \geq c_2\|v\|_X^{p-1} - \frac{\|k_2(v)\|_{L^1(Q_T)}}{\|v\|_X} =$$
$$\|v\|_X^{p-1}\left[c_2 - \frac{\|k_2(v)\|_{L^1(Q_T)}}{\|v\|_X^p}\right] \to +\infty$$

as $\|v\|_X \to \infty$ which implies coercivity of A. ∎

REMARK 1. According to part c) of the proof, if assumptions I–V are satisfied such that (1.3) holds in the following weaker form:

$$\sum_{i=0}^{n} a_i(t,x,\zeta_0,\zeta;v)\zeta_i \geq c_2|\zeta|^p - [k_2(v)](t,x) \tag{1.27}$$

then (1.8), (1.9) imply that $(u_k) \to u$ and $(Du_k) \to Du$ a.e. in Q_T and $(A(u_k)) \to A(u)$ weakly in $X^\star$ for a suitable subsequence.

2. Existence and uniqueness with discontinuous terms

In this section we shall consider equations (0.2) if the functions a_i have the following form:

$$\begin{aligned}(2.28)\quad & a_i(t,x,\zeta_0,\zeta;v) = \\ & b([H(v)](t,x))a_i^1(t,x,\zeta_0,\zeta) + \tilde{b}([G(v)](t,x))a_i^2(t,x,\zeta_0,\zeta),\end{aligned}$$

$i = 1,\dots,n$;

$$\begin{aligned}(2.29)\quad a_0(t,x,\zeta_0,\zeta;v) = {} & b_1([G_1(v)](t,x))a_0^1(t,x,\zeta_0,\zeta) + \\ & b_2([G_2(v)](t,x))a_0^2(t,x,\zeta_0,\zeta) + a_0^3(t,x,\zeta_0,\zeta)\end{aligned}$$

where a_i^1, a_i^2 $(i = 0,1,\dots,n)$ satisfy the Carathéodory conditions, a_0^3 is measurable;

$$(2.30)\qquad |a_i^1(t,x,\zeta_0,\zeta)| \le c_1(|\zeta_0|^{p-1} + |\zeta|^{p-1}) + k_1(x)$$

with some constant c_1, $k_1 \in L^q(\Omega)$, $i = 0,1,\dots,n$;

$$\begin{aligned}(2.31)\qquad & |a_i^2(t,x,\zeta_0,\zeta)| \le c_1(|\zeta_0|^\rho + |\zeta|^\rho), \\ & \qquad\qquad \text{with some } 0 \le \rho < p-1, \quad i = 0,1,\dots,n;\end{aligned}$$

$$(2.32)\qquad |a_i^3(t,x,\zeta_0,\zeta)| \le c_1|\zeta_0|^{p-1} + k_1(x).$$

Further,

$$(2.33)\qquad \sum_{i=1}^n [a_i^1(t,x,\zeta_0,\zeta) - a_i^1(t,x,\zeta_0,\zeta^\star)](\zeta_i - \zeta_i^\star) > 0 \text{ if } \zeta \ne \zeta^\star;$$

$$(2.34)\qquad \sum_{i=1}^n [a_i^2(t,x,\zeta_0,\zeta) - a_i^2(t,x,\zeta_0,\zeta^\star)](\zeta_i - \zeta_i^\star) \ge 0;$$

$$(2.35)\qquad \sum_{i=1}^n a_i^1(t,x,\zeta_0,\zeta)\zeta_i \ge c_2|\zeta|^p - k_2(x)$$

with some constant $c_2 > 0$, $k_2 \in L^1(\Omega)$;

$$(2.36)\qquad \sum_{i=1}^n a_i^2(t,x,\zeta_0,\zeta)\zeta_i \ge 0;$$

$$(2.37)\qquad a_0^1(t,x,\zeta_0,\zeta)\zeta_0 \ge c_2|\zeta_0|^p - k_2(x);$$

$$(2.38)\qquad a_0^3(t,x,\zeta_0,\zeta)\zeta_0 \ge 0.$$

The functions $b, \tilde{b} : R \to R$ are continuous, $b_i : R \to R$ are measurable and satisfy the following conditions: there exist positive constants c_3, c_4, c_5 such that

$$c_3 \le b(\theta) \le c_4, \qquad c_3 \le b_1(\theta) \le c_4, \tag{2.39}$$

$$0 \le \tilde{b}(\theta) \le c_5|\theta|^{p-1-\rho}, \qquad |b_2(\theta)| \le c_5|\theta|^{p-1-\rho_0} \tag{2.40}$$

with some $\rho_0 > \rho$.

Finally, the operators H, G, G_i satisfy:

$$H, G_1 : L^p(0, T; W^{1-\delta,p}(\Omega)) \to L^1(Q_T) \text{ and} \tag{2.41}$$

$$G, G_2 : L^p(0, T; W^{1-\delta,p}(\Omega)) \to L^p(Q_T) \tag{2.42}$$

are linear and continuous operators.

Since the functions b_i are locally bounded, we may define for any $\varepsilon > 0$ (see [12])

$$\overline{b}_i^{\varepsilon}(\theta) = \operatorname{ess\,sup}_{|\theta-\tilde{\theta}|<\varepsilon} b_i(\tilde{\theta}), \quad \underline{b}_i^{\varepsilon}(\theta) = \operatorname{ess\,inf}_{|\theta-\tilde{\theta}|<\varepsilon} b_i(\tilde{\theta}),$$

$$\overline{b}_i(\theta) = \lim_{\varepsilon\to 0} \overline{b}_i^{\varepsilon}(\theta), \quad \underline{b}_i(\theta) = \lim_{\varepsilon\to 0} \underline{b}_i^{\varepsilon}(\theta).$$

Similarly define

$$\overline{a}_0^{3,\varepsilon}(t,x,\zeta_0,\zeta) = \operatorname{ess\,sup}_{|\zeta_0-\tilde{\zeta}_0|<\varepsilon, |\zeta-\tilde{\zeta}|<\varepsilon} a_0^3(t,x,\tilde{\zeta}_0,\tilde{\zeta}),$$

$$\underline{a}_0^{3,\varepsilon}(t,x,\zeta_0,\zeta) = \operatorname{ess\,inf}_{|\zeta_0-\tilde{\zeta}_0|<\varepsilon, |\zeta-\tilde{\zeta}|<\varepsilon} a_0^3(t,x,\tilde{\zeta}_0,\tilde{\zeta}),$$

$$\overline{a}_0^3(t,x,\zeta_0,\zeta) = \lim_{\varepsilon\to 0} \overline{a}_0^{3,\varepsilon}(t,x,\zeta_0,\zeta),$$

$$\underline{a}_0^3(t,x,\zeta_0,\zeta) = \lim_{\varepsilon\to 0} \underline{a}_0^{3,\varepsilon}(t,x,\zeta_0,\zeta).$$

THEOREM 2.1. *Assume* (2.28)–(2.42). *Then for any* $f \in L^q(0, T; V^\star)$ *there exists* $u \in L^p(0, T; V)$, $\beta_1 \in L^\infty(Q_T)$, $\beta_2 \in L^{\frac{p}{p-1-\rho_0}}(Q_T)$, $\beta_3 \in$ $\in L^q(Q_T)$ *such that*

$$D_t u \in L^q(0, T; V^\star), \quad u(0) = 0, \tag{2.43}$$

$$[D_t u, v] + \sum_{i=1}^{n} \int_{Q_T} \Big\{ b([H(u)](t,x)) a_i^1(t,x,u,Du) +$$

$$+ \tilde{b}([G(u)](t,x)) a_i^2(t,x,u,Du) \Big\} D_i v \, dt dx + \tag{2.44}$$

$$+\int_{Q_T}\Big\{\beta_1(t,x)a_0^1(t,x,u,Du)+\beta_2(t,x)a_0^2(t,x,u,Du)+\beta_3(t,x)\Big\}\,v\,dt\,dx=[f,v]$$

and for a.e. $(t,x)\in Q_T$

$$\underline{b}_i([G_i(u)](t,x))\le\beta_i(t,x)\le\overline{b}_i([G_i(u)](t,x)),\quad i=1,2, \tag{2.45}$$
$$\underline{a}_0^3(t,x,u(t,x),Du(t,x))\le\beta_3(t,x)\le\overline{a}_0^3(t,x,u(t,x),Du(t,x)).$$

REMARK 2. Clearly, if b_i, a_0^3 are Carathéodory functions, (2.42)–(2.45) means that u is a weak solution in usual sense.

REMARK 3. The value of the operators H, G, G_i in (t,x) may have e.g. one of the following forms:

$u(\tau(t),x)$ where $\tau:[0,T]\to[0,T]$ is a C^1 function, $\tau(t)\le t,\quad \tau'(t)>0$;

$$\int_\Omega g_1(t,x,\xi)u(t,\xi)d\xi;\quad \int_{Q_t} g_2(t,\tau,x,\xi)u(\tau,\xi)d\xi d\tau;\quad \int_0^t g_3(t,\tau,x)u(\tau,x)d\tau;$$

$$\int_{\partial\Omega} g_4(t,x,\xi)u(t,\xi)d\sigma_\xi,\ \int_{\Gamma_t} g_5(t,\tau,x,\xi)u(\tau,\xi)d\tau d\sigma_\xi,\ \text{where } \Gamma_t=[0,t]\times\partial\Omega,$$

$$\text{ess sup}_{(t,x)\in Q_T}\int_\Omega |g_1(t,x,\xi)|^q d\xi<\infty,$$
$$\text{ess sup}_{(t,x)\in Q_T}\int_{Q_T} |g_2(t,\tau,x,\xi)|^q d\tau d\xi<\infty,$$
$$\text{ess sup}_{(t,x)\in Q_T}\int_0^t |g_3(t,\tau,x)|^q d\tau<\infty,$$
$$\text{ess sup}_{(t,x)\in Q_T}\int_{\partial\Omega} |g_4(t,\tau,\xi)|^q d\sigma_\xi<\infty,$$
$$\text{ess sup}_{(t,x)\in Q_T}\int_{\Gamma_T} |g_5(t,\tau,x,\xi)|^q d\tau d\sigma_\xi<\infty,\quad \delta<1-1/p.$$

PROOF OF THEOREM 2.1. Let functions $j\in C_0^\infty(R)$, $\tilde{j}\in C_0^\infty(R^{n+1})$ be supported by the unit ball with the properties

$$j,\tilde{j}\ge 0,\quad \int j=1,\quad \int\tilde{j}=1$$

and for any positive integer k define functions $j_k, \tilde{j}_k$ by

$$j_k(\theta) = kj(k\theta), \quad \tilde{j}_k(\zeta_0, \zeta) = k^{n+1}\tilde{j}(k\zeta_0, k\zeta).$$

Then the convolutions (with fixed t, x)

$$b_1^k = b_1 \star j_k, \quad b_2^k = b_2 \star j_k, \quad a_0^{3,k} = a_0^3 \star \tilde{j}_k \tag{2.46}$$

are smooth functions (of $\theta, (\zeta_0, \zeta)$, respectively). Then we may define operators

$$A, B_1^k, B_2^k, B_3^k : L^p(0, T; V) \to L^q(0, T; V^\star)$$

by

$$[A(u), v] = [A(u), v]_T =$$
$$= \sum_{i=1}^n \int_{Q_T} \{b([H(u)](t,x))a_i^1(t,x,u,Du) + \tilde{b}([G(u)](t,x))a_i^2(t,x,u,Du)\} D_i v \, dt dx,$$

$$[B_l^k(u), v] = [B_l^k(u), v]_T = \int_{Q_T} b_l^k([G_l(u)](t,x) a_0^l(t,x,u,Du) v \, dt dx, \; l = 1, 2,$$

$$[B_3^k(u), v] = [B_3^k(u), v]_T = \int_{Q_T} a_0^{3,k}(t,x,u,Du) v \, dt dx.$$

By Theorem 1.1 the operator

$$A + B_1^k + B_2^k + B_3^k : L^p(0, T; V) \to L^q(0, T; V^\star)$$

is bounded, demicontinuous, pseudomonotone with respect to $D(L)$ and it is coercive, because we can show that the functions a_i, $i = 1, \dots, n$, defined by (2.28) and $a_0 = a_0^k$, defined by

$$a_0^k(t,x,\zeta_0,\zeta; v) = b_1^k([G_1(v)](t,x)) a_0^1(t,x,\zeta_0,\zeta) +$$
$$+ b_2^k([G_2(v)](t,x)) a_0^2(t,x,\zeta_0,\zeta) + a_0^{3,k}(t,x,\zeta_0,\zeta)$$

satisfy the conditions I–V.

The conditions I, II follow easily from (2.30)–(2.32), (2.39)–(2.42) and Young's inequality since according to the definitions (2.46)

$$c_3 \le b_1^k(\theta) \le c_4, \tag{2.47}$$

$$|b_2^k(\theta)| \le c_5'|\theta|^{p-1-\rho_0} + c_6', \quad |a_0^{3,k}(t,x,\zeta_0,\zeta)| \le c_1'|\zeta_0|^{p-1} + \tilde{k}_1(x) \tag{2.48}$$

with some constants c_1', c_5', c_6' and $\tilde{k}_1 \in L^q(\Omega)$.

The condition III follows directly from (2.33), (2.34), (2.39). Further, condition IV follows from (2.31), (2.32), (2.35)–(2.38), (2.42), (2.47), (2.48) since by (2.32), (2.38) and (2.46)

$$a_0^{3,k}(t,x,\zeta_0,\zeta)\zeta_0 \geq -k^\star(x)$$

with some $k^\star \in L^1(\Omega)$ and from (2.31), (2.42), (2.48) we obtain by using Young's inequality for any $\varepsilon > 0$

$$|b_2^k([G_2(v)](t,x))a_i^2(t,x,\zeta_0,\zeta)\zeta_0| \leq$$
$$\leq \varepsilon |a_i^2(t,x,\zeta_0,\zeta)\zeta_0|^{\frac{p}{\rho+1}} + c(\varepsilon)|b_2^k([G_2(v)](t,x))|^{\frac{p}{p-1-\rho}} \leq$$
$$\leq c_1^\star \varepsilon(|\zeta_0|^\rho + |\zeta|^\rho) + c^\star(\varepsilon)\left\{|[G_2(v)](t,x)|^{p\frac{p-1-\rho_0}{p-1-\rho}} + 1\right\}$$

with some constant $c_1^\star$ (not depending on ε) and a constant $c^\star(\varepsilon)$ (depending on ε). Choosing sufficiently small $\varepsilon > 0$, we obtain IV with $c_2/2$ instead of c_2 and

$$[k_2(v)](t,x) = c^\star(\varepsilon)\left\{|[G_2(v)](t,x)|^{p\frac{p-1-\rho_0}{p-1-\rho}} + 1\right\}$$

by (2.42). ($k_2(v)$ is independent of k.)

Finally, condition V follows from (2.41), (2.42) and the fact that a_i^1, a_i^2 ($i = 0, 1, \ldots, n$), $a_0^{3,k}$ satisfy the Carathéodory condition; $b, \tilde{b}$ are continuous, b_1^k, b_2^k are smooth functions.
Because, if $(u_k) \to$ strongly in $L^p(0,T; W^{1-\delta,p}(\Omega))$ then by (2.41), (2.42)

$$(H(u_k)) \to H(u), \quad (G_1(u_k)) \to G_1(u) \tag{2.49}$$

in $L^1(Q_T)$ and

$$(G(u_k)) \to G(u), \quad (G_2(u_k)) \to G_2(u) \tag{2.50}$$

in $L^p(Q_T)$, consequently, (2.49), (2.50) hold a.e. in Q_T for a subsequence.

So we have shown that all the conditions of Theorem 1.1 are satisfied, thus for each k there exists

$$u_k \in L^p(0,T; V) \text{ with } D_t u_k \in L^q(0,T; V^\star), \quad u_k(0) = 0, \tag{2.51}$$

$$D_t u_k + (A + B_1^k + B_2^k + B_3^k)(u_k) = f. \tag{2.52}$$

Consequently, for any $t \in [0, T]$

$$\begin{aligned}(2.53) \qquad & \int_0^t \langle D_t u_k(\tau), u_k(\tau)\rangle d\tau + [A(u_k), u_k]_t + [B_1^k(u_k), u_k]_t + \\ & + [B_2^k(u_k), u_k]_t + [B_3^k(u_k), u_k]_t = [f, u_k]_t.\end{aligned}$$

Since IV is satisfied with a constant c_2 and $k_2(v)$ which are independent of k, we obtain from (2.53) that

$$(2.54) \qquad \|u_k\|_X, \quad \|(A + B_1^k + B_2^k + B_3^k)(u_k)\|_{X^\star} \text{ are bounded,}$$

consequently,

$$(2.55) \qquad \|D_t u_k\|_{X^\star} \text{ is bounded,}$$

$$(2.56) \qquad \|u_k\|_{L^\infty(0,T;L^2(\Omega))} \text{ is bounded.}$$

(2.54), (2.55) imply that

$$\begin{aligned}(2.57) \quad & (u_k) \to u \text{ weakly in } X, \text{ strongly in } L^p(0, T; W^{1-\delta,p}(\Omega)) \text{ and} \\ & (D_t u_k) \to D_t u \text{ weakly in } X^\star\end{aligned}$$

for a subsequence (see, e.g., [11]).

Since $(u_k) \to u$ in $L^p(Q_T)$, it is not difficult to show (by using the assumptions of our theorem and Hölder's inequality) that

$$(2.58) \qquad \lim_{k\to\infty} [B_i^k(u_k), u_k - u] = 0 \quad (i = 1, 2, 3), \quad \lim_{k\to\infty} [f, u_k - u] = 0.$$

Thus, applying (2.52) to $(u_k - u)$, we find

$$(2.59) \qquad \lim_{k\to\infty} [A(u_k), u_k - u] = 0.$$

Now we apply Remark 1 to operator A which satisfies (1.27), so by (2.57), (2.59) we obtain that

$$\begin{aligned}(2.60) \quad & (u_k) \to u, \quad (D_t u_k) \to D_t u \\ & \text{a.e. in } Q_T \text{ and } (A(u_k)) \to A(u) \text{ weakly in } X^\star\end{aligned}$$

for a subsequence.

By (2.47) $(b_1^k([G_1(u_k)](t, x)))$ is a bounded sequence in $L^\infty(Q_T)$, i.e. it is a bounded sequence of linear continuous functionals on $L^1(Q_T)$, thus there exists $\beta_1 \in L^\infty(Q_T)$ such that for a subsequence

$$(2.61) \qquad \lim_{k\to\infty} \int_{Q_T} b_1^k(G_1(u_k)) g\, dt\, dx = \int_{Q_T} \beta_1 g\, dt\, dx \text{ for any } g \in L^1(Q_T).$$

Consequently, for any $v \in X$

$$\int_{Q_T} b_1(G_1(u_k))a_0^1(t,x,u_k,Du_k)v\,dt\,dx = \int_{Q_T} b_1(G_1(u_k))a_0^1(t,x,u,Du)v\,dt\,dx +$$

$$+\int_{Q_T} b_1(G_1(u_k))[a_0^1(t,x,u_k,Du_k) - a_0^1(t,x,u,Du)]v\,dt\,dx \to \int_{Q_T} \beta_1 v\,dt\,dx$$

because the second term in the right hand side of the equality tends to 0 by (2.47), (2.60) and Vitali's theorem. Thus

$$(B_1^k(u_k)) \to \beta_1 a_0^1(t,x,u,Du) \text{ weakly in } X^\star. \tag{2.62}$$

Similarly, $a_0^{3,k}(t,x,u_k,Du_k)$ is bounded in $L^q(Q_T)$, hence there exists $\beta_3 \in L^q(Q_T)$ such that (for a subsequence)

$$a_0^{3,k}(t,x,u_k,Du_k) \to \beta_3 \text{ weakly in } L^q(Q_T) \text{ as } k \to \infty. \tag{2.63}$$

Finally,

$$(b_2^k(G_2(u_k))) \to \beta_2 \text{ weakly in } L^{\frac{p}{p-1-\varrho_0}}(Q_T), \tag{2.64}$$

hence, by using Hölder's inequality, we obtain that

$$(B_2^k(u_k)) \to \beta_2 a_0^2(t,x,u,Du) \text{ weakly in } X^\star \tag{2.65}$$

with some $\beta_2 \in L^{\frac{p}{p-1-\varrho_0}}(Q_T)$. By using (2.60)–(2.65), we obtain from (2.52) as $k \to \infty$ the equality (2.44).

To complete the proof of our theorem, we have to show the inequalities (2.45). Now we use arguments of [12]. Because of (2.57), (2.41), (2.42)

$$G_1(u_k) \to G_1(u), \quad G_2(u_k) \to G_2(u) \tag{2.66}$$

in $L^1(Q_T)$ and so for a subsequence (2.66) holds a.e. in Q_T. Thus for arbitrary small number $a > 0$ there exists a set $\omega \subset Q_T$ such that

$$\begin{aligned} &\lambda(\omega) < a \text{ and } G_i(u_k) \to G_i(u) \text{ uniformly in } Q_T \setminus \omega, \\ &\qquad G_i(u) \in L^\infty(Q_T \setminus \omega), \quad (i = 1,2). \end{aligned} \tag{2.67}$$

Hence for any $\varepsilon > 0$ there exists $k_0 > 2/\varepsilon$ such that

$$|[G_i(u_k)](t,x) - [G_i(u)](t,x)| < \varepsilon/2 \text{ if } (t,x) \in Q_T \setminus \omega, \quad k > k_0. \tag{2.68}$$

By definition $b_i^k([G_i(u_k)](t,x)$ is an average of b_i over the interval centered in $[G_i(u_k)](t,x)$ with radius $1/k < \varepsilon/2$ and this interval is contained by the interval centered in $[G_i(u)](t,x)$ with radius ε, because of (2.68). Thus

$$\underline{b}_i^\varepsilon([G_i(u)](t,x)) \le b_i^k([G_i(u_k)](t,x)) \le \overline{b}_i^\varepsilon([G_i(u)](t,x)).$$

Hence for any $\varphi \in C_0^\infty(Q_T)$, $\varphi \ge 0$

$$\int_{Q_T\setminus\omega} \underline{b}_i^\varepsilon([G_i(u)](t,x))\varphi(t,x)dtdx \le \int_{Q_T\setminus\omega} b_i^k([G_i(u_k)](t,x))\varphi(t,x)dtdx \le$$
$$\le \int_{Q_T\setminus\omega} \overline{b}_i^\varepsilon([G_i(u)](t,x))\varphi(t,x)dtdx$$

which implies by (2.61), (2.64)

$$\int_{Q_T\setminus\omega} \underline{b}_i^\varepsilon([G_i(u)](t,x))\varphi(t,x)dtdx \le \int_{Q_T\setminus\omega} \beta_i(t,x)\varphi(t,x)dtdx \le \tag{2.69}$$
$$\le \int_{Q_T\setminus\omega} \overline{b}_i^\varepsilon([G_i(u)](t,x))\varphi(t,x)dtdx.$$

The functions $G_i(u)$ $(i = 1,2)$ are bounded in $Q_T \setminus \omega$ thus we obtain from (2.69) as $\varepsilon \to 0$

$$\int_{Q_T\setminus\omega} \underline{b}_i([G_i(u)](t,x))\varphi(t,x)dtdx \le \int_{Q_T\setminus\omega} \beta_i(t,x)\varphi(t,x)dtdx$$
$$\le \int_{Q_T\setminus\omega} \overline{b}_i([G_i(u)](t,x))\varphi(t,x)dtdx.$$

Since $\varphi \in C_0^\infty(Q_T)$ is an arbitrary nonnegative function, we find

$$\underline{b}_i([G_i(u)](t,x)) \le \beta_i(t,x) \le \overline{b}_i([G_i(u)](t,x)), \quad i = 1,2 \tag{2.70}$$

for a.e. $(t,x) \in Q_T \setminus \omega$. The equality (2.70) holds for arbitrary $a > 0$, $\lambda(\omega) < a$, $\omega \subset Q_T$, thus it is valid a.e. in Q_T.

Similarly can be proved the inequality for $\beta_3(t,x)$, by using (2.60). ∎

If certain additional conditions are satisfied then one can prove uniqueness of the solution.

THEOREM 2.2. *Let the assumptions of Theorem 2.1 be satisfied such that*

$$b(\theta) = \tilde{b}(\theta) = b_1(\theta) = 1 \text{ for all } \theta \in R,$$
$$a_i^1(t,x,\zeta_0,\zeta) = \alpha_i^1(t,x,\zeta), \quad i = 1,\dots,n$$

(it is not depending on ζ_0),

$$a_0^1(t,x,\zeta_0,\zeta) = \alpha_0^1(t,x,\zeta_0)$$

(it is not depending on ζ) and the function $\zeta_0 \to \alpha_0^1(t,x,\zeta_0)$ is monotone nondecreasing for a.e. (t,x). Further, $a_0^2(t,x,\zeta_0,\zeta) = 1$, b_2 satisfies a global Lipschitz condition and G_2 is continuous with respect to the norm of $L^2(Q_T)$ such that

$$\|\mathrm{e}^{-c_0 t}[G_2(\mathrm{e}^{c_0\tau}\tilde{u})]\|_{L^2(Q_T)} \le const\|\tilde{u}\|_{L^2(Q_T)}$$

for any constant $c_0 > 0$ where the constant is not depending on $c_0, \tilde{u}$. ($\mathrm{e}^{c_0\tau}\tilde{u}$ denotes the function $(\tau,\xi) \mapsto \mathrm{e}^{c_0\tau}\tilde{u}(\tau,\xi)$.) Finally,

$$a_0^3(t,x,\zeta_0,\zeta) = \alpha_0^3(\zeta_0)$$

where α_0^3 is monotone nondecreasing.

Then the solution of (2.43)–(2.45) *is unique.*

PROOF. Assume that u and $u^\bigstar$ are solutions of (2.43)–(2.45). Define $\tilde{u}$ and $\tilde{u}^\bigstar$ by

$$\tilde{u}(t,x) = \mathrm{e}^{-c_0 t}u(t,x), \quad \tilde{u}^\bigstar(t,x) = \mathrm{e}^{-c_0 t}u^\bigstar(t,x)$$

with some (sufficiently large) constant $c_0 > 0$. Transforming (2.44) to $\tilde{u}$ and $\tilde{u}^\bigstar$ and applying them to $v = \tilde{u} - \tilde{u}^\bigstar$, we obtain

$$[D_t(\tilde{u} - \tilde{u}^\bigstar), \tilde{u} - \tilde{u}^\bigstar] + c_0 \int_{Q_T} (\tilde{u} - \tilde{u}^\bigstar)^2 dt dx + \tag{2.71}$$

$$+ \int_{Q_T} \mathrm{e}^{-2c_0 t} \sum_{i=1}^n \left[\alpha_i^1(t,x,\mathrm{e}^{c_0 t}D\tilde{u}) - \alpha_i^1(t,x,\mathrm{e}^{c_0 t}D\tilde{u}^\bigstar)\right](\mathrm{e}^{c_0 t}D_i\tilde{u} - \mathrm{e}^{c_0 t}D_i\tilde{u}^\bigstar) dt dx +$$

$$+ \int_{Q_T} \mathrm{e}^{-2c_0 t} \sum_{i=1}^n \left[\alpha_i^2(t,x,\mathrm{e}^{c_0 t}D\tilde{u}) - \alpha_i^2(t,x,\mathrm{e}^{c_0 t}D\tilde{u}^\bigstar)\right](\mathrm{e}^{c_0 t}D_i\tilde{u} - \mathrm{e}^{c_0 t}D_i\tilde{u}^\bigstar) dt dx +$$

$$+\int_{Q_T} e^{-2c_0 t}\left[\alpha_0^1(t,x,e^{c_0 t}\tilde{u}) - \alpha_0^1(t,x,e^{c_0 t}\tilde{u}^\star)\right](e^{c_0 t}\tilde{u} - e^{c_0 t}\tilde{u}^\star)dtdx+$$

$$+\int_{Q_T} e^{-2c_0 t}\left[\beta_2(t,x) - \beta_2^\star(t,x)\right](e^{c_0 t}\tilde{u} - e^{c_0 t}\tilde{u}^\star)dtdx+$$

$$+\int_{Q_T} e^{-2c_0 t}\left[\beta_3(t,x) - \beta_3^\star(t,x)\right](e^{c_0 t}\tilde{u} - e^{c_0 t}\tilde{u}^\star)dtdx = 0.$$

By the assumptions of our theorem, several terms in (2.71) are nonnegative, so we obtain from (2.71)

$$(2.72)\ c_0\int_{Q_T}(\tilde{u} - \tilde{u}^\star)^2 dtdx + \int_{Q_T} e^{-2c_0 t}[\beta_2(t,x) - \beta_2^\star(t,x)](u - u^\star)dtdx +$$
$$+\int_{Q_T} e^{-2c_0 t}[\beta_3(t,x) - \beta_3^\star(t,x)](u - u^\star)dtdx \le 0.$$

Since α_0^3 is monotone nondecreasing and

$$\lim_{\theta\to u(t,x)-0}\alpha_0^3(\theta) = \underline{\alpha}_0^3(u(t,x)) \le \beta_3(t,x) \le \overline{\alpha}_0^3(u(t,x)) = \lim_{\theta\to u(t,x)+0}\alpha_0^3(\theta)$$

and, similarly,

$$\lim_{\theta\to u^\star(t,x)-0}\alpha_0^3(\theta) \le \beta_3^\star(t,x) \le \lim_{\theta\to u^\star(t,x)+0}\alpha_0^3(\theta),$$

we obtain that $u(t,x) > u^\star(t,x)$ implies

$$\beta_3^\star(t,x) \le \lim_{\theta\to u^\star(t,x)+0}\alpha_0^3(\theta) \le \lim_{\theta\to u(t,x)-0}\alpha_0^3(\theta) \le \beta_3(t,x)$$

and $u(t,x) < u^\star(t,x)$ implies $\beta_3(t,x) \le \beta_3^\star(t,x)$, consequently, the third term in (2.72) is nonnegative.

Further, since b_2 is (globally) Lipschitz continuous, for the second term in (2.72) we have the estimate

$$\left|\int_{Q_T} e^{-2c_0 t}[\beta_2(t,x) - \beta_2^\star(t,x)](u - u^\star)dtdx\right| =$$

$$= \left| \int_{Q_T} e^{-c_0 t} \{ b_2([G_2(u)](t,x)) - b_2([G_2(u^\star)](t,x)) \} (\tilde{u} - \tilde{u}^\star) dt dx \right| \le$$

$$\le \tilde{c} \| e^{-c_0 t} [G_2(u) - G_2(u^\star)] \|_{L^2(Q_T)} \| \tilde{u} - \tilde{u}^\star \|_{L^2(Q_T)} \le \hat{c} \| \tilde{u} - \tilde{u}^\star \|^2_{L^2(Q_T)}$$

with some constants $\tilde{c}, \hat{c}$, not depending on c_0. Thus, choosing sufficiently large c_0, from (2.72) we obtain

$$\| \tilde{u} - \tilde{u}^\star \|^2_{L^2(Q_T)} = 0, \text{ i.e. } u = u^\star \text{ a.e.} \qquad \blacksquare$$

References

[1] R. A. ADAMS, *Sobolev spaces*, Academic Press, New York–San Francisco–London, 1975.

[2] J. BERKOVITS and V. MUSTONEN, Topological degreee for perturbations of linear maximal monotone mappings and applications to a class of parabolic problems, *Rend. Mat., Ser. VII* **12** (1992), Roma, 597–621.

[3] F. E. BROWDER, Pseudo-monotone operators and nonlinear elliptic boundary value problems on unbounded domains, *Proc. Natl. Acad. Sci. USA* **74** (1977), 2659–2661.

[4] M. CHIPOT and L. MOLINET, Asymptotic behavior of some nonlocal diffusion problems, *Applicable Analysis* **80** (2001), 279–315.

[5] M. CHIPOT and B. LOVAT, Existence and uniqueness results for a class of nonlocal elliptic and parabolic problems, advances in quenching; *Dyn. Contin. Discrete Impuls. Syst., Ser. A Math. Anal.* **8(1)** (2001), 35–51.

[6] J. I. DÍAZ and G. HETZER, A quasilinear functional reaction-diffusion equation arising in climatology, in: *PDEs and applications*, Dunod, Paris, 1998.

[7] J. I. DÍAZ and F. DE THELIN, On a nonlinear parabolic problem arising in some models related to turbulent flows, *SIAM J. Math. Anal.* **25** (1994), 1085–1111.

[8] J. I. DÍAZ, J. HERNÁNDEZ and L. TELLO, On the multiplicity of equilibrium solutions to a nonlinear diffusion equation on a manifold arising in climatology, *J. Math. Anal. Appl.* **216** (1997), 593–613.

[9] U. HORNUNG and W. JÄGER, Diffusion, absorption and reaction of chemicals in porous media; *J. Differ. Equat.* **92** (1991), 199–225.

[10] U. HORNUNG, W. JÄGER and A. MIKELIC, Reactive transport through an array of cells with semi-permeable membranes, *Math. Modelling Num. Anal.* **28** (1994), 59–94.

[11] J. L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[12] J. RAUCH, Discontinuous semilinear differential equations and multiple valued maps, *Proc. Am. Math. Soc.* **64** (1977), 277–282.

[13] L. SIMON, On the stabilization of solutions of nonlinear parabolic functional differential equations, *Proceedings of the Conference FSDONA-99, Syöte, 1999*, 239–250.

[14] E. ZEIDLER, *Nonlinear functional analysis and its applications II A and II B*, Springer, 1990.

László Simon

Department of Applied Analysis,
Loránd Eötvös University,
Pázmány P. sétány 1./c
H-1117 Budapest, Hungary
simonl@ludens.elte.hu

INDEX

ISSN 0524-9007

Address:
MATHEMATICAL INSTITUTE, EÖTVÖS LORÁND UNIVERSITY
ELTE, TTK, KARI KÖNYVTÁR, MATEMATIKAI TUDOMÁNYÁGI SZAKGYŰJTEMÉNY
PÁZMÁNY PÉTER SÉTÁNY 1/c, 1117 BUDAPEST, HUNGARY
Műszaki szerkesztő:

Fried Katalin PhD

A kiadásért felelős: az Eötvös Loránd Tudományegyetem rektora
A kézirat a nyomdába érkezett: 2005 április. Megjelent: 2005. május
Terjedelem: 17,16 A/4 ív. Példányszám: 500
Készült az EMTEX szedőprogram felhasználásával
az MSZ 5601–59 és 5602–55 szabványok szerint
Az elektronikus tipografálás Bori Tamás munkája
Nyomda: Haxel Bt.
Felelős vezető: Korándi József