MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 16/1976.

TETSZŐLEGESEN NAGY EGÉSZ SZÁMOKKAL VALÓ PONTOS SZÁMOLÁS SZÁMITÓGÉPPEL

Gesztelyi Ernő – Jékel Pál
(Debrecen, KLTE Számitástudományi Tanszék)

0. Bevezetés

Adott számitógéppel rendelkező számolóközpont kerülhet olyan feladat elé, hogy olyan nagy egész számokkal kellene aritmetikai müveleteket pontosan végezni, amelyek nagyobbak mint a gépen ábrázolható legnagyobb egész szám. Természetesen sok esetben megoldható a probléma, ha dupla, tripla, stb. hosszuságu számokkal dolgozunk, az ALGOL 68 terminológiájával kifejezve: ha rátérünk a *long integer*, vagy *long long integer*, stb. módu számokkal való számolásra. Természetesen adott gépi reprezentáció esetén a *long*-ok száma korlátozott, és ezért adott gépen igy nem minden feladat végezhető el. (Tudomásunk szerint nagy teljesitményü számitógépekkel végeznek olyan vizsgálatokat, amelyekkel igen nagy számokra vonatkozó, számelméleti érdekességgel biró kérdéseket döntenek el.)

Jelen dolgozatnak az a célja, hogy megmutassunk egy módszert arra, hogy hogyan lehet bármilyen nagy egész számokkal pontos számitásokat végeztetni rekeszek összekapcsolása nélkül. Természetesen módszerünk akkor is alkalmazható, ha mód van rekeszek összekapcsolására, ebben az esetben a programok futási ideje rövidül le.

A módszer alapgondolata a következő: Tekintsünk valamilyen b alapot, és az egész számokat b alapú számrendszerben megadva képzeljük el. Akkor adott szám esetén minden számjegyet elhelyezhetünk egy egy memória rekeszben. Ilyen módon igen nagy egész számok ábrázolhatók a gépben, tekintve, hogy még az operativ memória rekeszeinek a száma is tekintélyes, nem is beszélve arról, hogy dobra is vihetők. Általában azonban nincs szükség a számoknak a memóriában való tárolására, konkrét esetekben elegendő a számjegyeket előállitó valamilyen eljárás algoritmusát tárolni a memóriában.

Legyen tehát $f(n)$ és $g(n)$ két függvényeljárás, melyek egy-egy pozitiv egész szám b alapu számjegyeit generálják ugy, hogy a függvényértékek a b^n együtthatóját szolgáltatják. (Megjegyezzük, hogy ha egy számot a fentebb leirt módon tároljuk a memóriában, akkor célszerü egy vektorban elhelyezni, és akkor (az ALGOL 60 terminológiáját használva) $f[n]$-et és $g[n]$-et irunk). Akkor felirhatunk egy eljárást, mely az $f(n) + g(n)$ értékeket generálja. Mivel általában $f(n) + g(n) > b$ ezért $f(n) + g(n)$ nem adja meg általában valamilyen egész szám b alapu számjegyeit. A dolgozatban megadunk egy Tb nevü eljárást, melynek alkalmazásával $Tb(f(n) + g(n))$ már olyan függvény lesz, amely megadja a

$$\sum_{n=0}^{N} (f(n) + g(n))b^n$$

szám b alapu számjegyeit, ahol N egy bizonyos szám ugy, hogy $f(n) = 0$ ha $n > N$.

Ha az f eljárás és a g eljárás által meghatározott egész számok szorzatát akarjuk megkapni akkor képezni kell a

$$\sum_{k=0}^{N} f(n-k)g(k)$$

konvolueiót, és azután alávetni a Tb eljárásnak, hogy megkapjuk a szorzat számjegyeit.

Ha rekeszeket kapcsolunk össze, akkor valamilyen módon (hardware, vagy software eszközökkel) meg kell oldani az átvitel próbjémáját. A mi esetünkben tehát a müveletek végzése és az átvitel külön van választva. Ha több müveletet kell egy feladat kapcsán elvégezni, akkor nem szükséges minden művelet elvégzése után alkalmazni a Tb eljárást. Ha túlcsordulás veszélye nem áll fenn, akkor elég csak a müveletsor végén alkalmazni a Tb eljárást, hogy a végeredményt megkapjuk.

Célszerü b-t minél nagyobbra választani, és nagyságát csak a túlcsordulás korlátozza. Ugyanis minél nagyobb a b annál jobban kihasználhatjuk a gép gyors aritmetikáját.

Gyakorlatilag célszerü az is, ha $b = 10^m$ (m pozitiv egész) mert ekkor nincs szükség olyan algoritmusra, mely a b alapu számrendszerből a 10-es alapu számrendszerbe alakit át. Ugyanis a kiiratásnál elegendő a rekeszek tartalmát a legnagyobb helyértékü tagtól kezdve visszafelé kiiratni egymás mellé, és akkor megkapjuk a szám 10-es számrendszerbeli alakját.

A dolgozat tárgyalásmódja a matematikai precizitás követelményeit igyekszik megvalósitani. Enélkül ugyanis nem lehetünk biztosak abban, hogy a gép valóban azt számolja-e ki, amit akarunk.

A dolgozat végén egy programot közlünk ALGOL 60 nyelven, amelyet az ODRA 1204 gépen le is futtattunk. A program táblázatot készit a pozitiv egész számok faktoriálisairól 1-től kezdve folyamatosan.

1. Egész értékeket felvevő véges tartószámú aritmetikai függvények

1.1 **Definició.** I-vel jelöljük azoknak az f függvényeknek a halmazát, melyek a következő tulajdonságoknak tesznek eleget.

a.) f értelmezési tartománya a nem negativ egész számok halmaza.
b.) Az f értékkészlete az egész számok halmazának részhalmaza.
c.) Minden $f \neq 0$ függvényhez f-től függően tartozik egy N nemnegativ egész úgy, hogy $f(N) \neq 0$, de ha $n > N$ akkor $f(n) = 0$.
d.) Az azonosan zérus értéket felvevő függvény I-be tartozik.

1.2 **Definició.** Legyen $f \in I$ és $f \neq 0$. Azt a N számot, amelyre $f(N) \neq 0$ de minden $n > N$ mellett $f(n) = 0$, az f függvény felső tartó számának nevezzük és Supp$f$-fel jelöljük, vagyis:

(1) $N = \text{Supp}f$, ha $f(N) \neq 0$, de minden $n > N$ mellett $f(n) = 0$.

1.3. **Definició.** *I-ben* értelmezzük az összeadást és a konvolúciós szorzást a következő módon: Legyen $f,g \in I$, akkor

(2) $(f + g)(n) = f(n) + g(n)$

(3) $(f * g)(n) = \sum_{k=0}^{n} f(n - k)g(k).$

1.1. **Tétel.** *Ha* $f,g \in I$ *akkor* $f + g \in I$ *és* $f * g \in I$ *és ha* $f,g,f + g \neq 0$ *akkor*

(4) $\text{Supp}(f + g) \leqslant \max(\text{Supp}f, \text{Supp}g)$ *és ha* $f,g \neq 0$ *akkor* $f * g \neq 0$ *és*

(5) $\text{Supp}(f * g) = \text{Supp}f + \text{Supp}g.$

Bizonyitás. Az 1.1. definició *a* és *b* tulajdonságának nyilván eleget tesz $f + g$ is és $f*g$ is, ha $f,g \in I$. Mivel az is triviális, hogy ha f és g közül legalább az egyik azonosan zérus, akkor $f + g$ is és $f*g$ is I-be tartozik, elég csak a (4) és (5) tulajdonságot igazolni, mert ezekből a c) tulajdonság az összegre és a konvolúcióra már következik.
Legyen tehát $f,g \in I$. Ha $f + g = 0$, akkor az 1. definició d) miatt $f + g \in I$.

Tegyük fel, hogy $f + g \neq 0$, és legyen $N = \max(\text{Supp}f, \text{Supp}g)$. Ha $n > N$, akkor igy $f(n) = 0$ és $g(n) = 0$, tehát

(6) $(f + g)(n) = f(n) + g(n) = 0 \qquad (n > N).$

Mivel $f + g \neq 0$, ezért van olyan legnagyobb N_0 melyre $(f + g)(N_0) \neq 0$ és (6) miatt nyilván $N_0 \leqslant N$.

Mivel igy $N_0 = \text{Supp}(f + g)$, ezért a (4) egyenlőtlenséget igazoltuk.

Most rátérünk az (5) egyenlőség igazolására, ahol feltesszük, hogy $f,g \neq 0$ és $f,g \in I$. $f,g \neq 0$ miatt létezik

(7) $N_1 = \text{Supp}f$

és

(8) $N_2 = \text{Supp}g.$

Először megmutatjuk, hogy ha

(9) $N_0 = N_1 + N_2$

akkor

(10) $(f * g)(N_0) = \sum_{k=0}^{N_0} f(N_0 - k)g(k) \neq 0,$

mert, mint igazolni fogjuk,

(11) $\quad (f * g)(N_0) = f(N_1)g(N_2),$

és mivel (7) és (8) miatt $f(N_1) \neq 0$ és $g(N_2) \neq 0$, igy $(f*g)(N_0) = f(N_1)g(N_2) \neq 0$. Ha $k = N_2$, akkor $f(N_0 - k)g(k) = f(N_1)g(N_2)$. Tehát (10) belátásához azt kell kimutatni, hogy ha $k \neq N_2$, akkor

(12) $\quad f(N_0 - k)g(k) = 0.$

Valóban, ha $k < N_2$, akkor $N_0 - k > N_1$ és igy (7) miatt $f(N_0 - k) = 0$ következtében teljesül (12). Ha pedig $k > N_2$, akkor (8) miatt $g(k) = 0$ és igy megint igaz a (12) egyenlőség. Tehát a (10) egyenlőség valóban fennáll.

Most azt igazoljuk, hogy ha $n > N_0$, akkor $(f * g)(n) = 0$. Valóban, ha a $\sum_{k=0}^{n} f(n-k)g(k)$ összegben $k \leqslant N_2$, akkor $n - k > N_1$ és akkor $f(n-k) = 0$, ha pedig $k > N_2$, akkor $g(k) = 0$, és igy $(f * g)(n) = 0$. Tehát $N_0 = \mathrm{Supp}(f * g)$ és igy az (5) egyenlőség (7), (8) és (9) következménye.

1.2. **Tétel.** *Az I halmaz a* (2) *és* (3) *alatt értelmezett műveletekre nézve nullosztómentes kommutativ gyűrűt alkot.**

Bizonyitás. Az 1.1. tételben igazoltuk, hogy a (2) és (3) alatti müveletek nem vezetnek ki az I-ből. Könnyen igazolható, hogy I az összeadásra nézve Ábel-csoportot alkot. Ismeretes, hogy a (3) alatti konvolúció kommutativ, associativ és az összeadásra nézve disztributiv művelet. Igy tehát I valóban kommutativ gyürü. A nullosztómentesség (5) következménye.

2. A szummációs transzformáció

2.1. **Definició.** Legyen $b > 1$ egész szám. I-ben értelmezünk egy S_b-vel jelölt transzformációt, melyet "b-re vonatkozó szummációs transzformációnak" nevezünk, és igy értelmezzük:

(1) $$(S_b f)(n) = \begin{cases} \sum_{k=0}^{\mathrm{Supp}f} f(k)b^k, & \text{ha } n = 0 \\ 0, & \text{ha } n > 0 \end{cases} \qquad (f \epsilon I)$$

Megjegyzés. Mivel $f(k) = 0$, ha $k > \mathrm{Supp}f$, ezért azt is irhatjuk, hogy

(2) $\quad (S_b f)(0) = \sum_{k=0}^{\infty} f(k)b^k.$

* L.BERG: EINFÜHRUNG IN DIE OPERATORENRECHNUNG, Berlin, 1962.

Megjegyzés. Világos, hogy ha $f \in I$, akkor $S_b f \in I$, azaz $S_b : I \longrightarrow I$.

Megjegyzés. Nyilvánvaló az is, hogy ha $(S_b f)(0) \neq 0$, akkor

(3) $\quad \mathrm{Supp}(S_b f) = 0.$

2.1.**Tétel.** *Az* S_b *transzformáció additiv*:

(4) $\quad S_b(f + g) = S_b f + S_b g \qquad (f, g \in I).$

Bizonyítás. Ha $n > 0$, nyilván igaz, hogy

(5) $\quad (S_b(f + g))(n) = (S_b f)(n) + (S_b g)(n),$

mert mindkét oldal zérus. Ha $n = 0$, akkor pedig

$$(S_b(f + g))(0) = \sum_{k=0}^{\infty} (f + g)(k) b^k = \sum_{k=0}^{\infty} f(k) b^k + \sum_{k=0}^{\infty} g(k) b^k =$$

$$= (S_b f)(0) + (S_b g)(0),$$

és igy (5) minden $n = 0,1,2, \ldots$ egészre teljesül, ami (4)-et igazolja.

2.2.**Tétel.** Az S_b transzformáció multiplikativ:

(6) $\quad S_b(f * g) = (S_b f) * (S_b g) \qquad (f, g \in I)$

és

(7) $\quad (S_b(f * g))(0) = (S_b f)(0) \cdot (S_b g)(0).$

Bizonyítás. Először a (7) egyenlőséget igazoljuk, vagyis azt, hogy

(8) $$\sum_{n=0}^{\mathrm{Supp}(f * g)} (f * g)(n) b^n = \sum_{n=0}^{\mathrm{Supp}\, f} f(n) b^n \cdot \sum_{n=0}^{\mathrm{Supp}\, g} g(n) b^n.$$

Legyen $N_1 = \mathrm{Supp} f$ és $N_2 = \mathrm{Supp} g$ (feltehetjük ugyanis, hogy $f, g \neq 0$, mert ha f és g egyike is azonosan zérus, akkor (6) és (7) nyilvánvalóan teljesül). Akkor $N_0 = N_1 + N_2$ jelöléssel

$$\sum_{n=0}^{\mathrm{Supp}(f*g)} (f*g)(n)b^n = \sum_{n=0}^{N_0}\sum_{k=0}^{n} f(n-k)g(k)b^n = \sum_{k=0}^{N_0}\sum_{n=k}^{N_0} f(n-k)g(k)b^n =$$

$$= \sum_{k=0}^{N_0} g(k) \sum_{n=k}^{N_0} f(n-k)b^n = \sum_{k=0}^{N_2} g(k) \sum_{m=0}^{N_0-k} f(m)b^{k+m} =$$

$$= \sum_{k=0}^{N_2} g(k)b^k \sum_{m=0}^{N_0-k} f(m)b^m = \sum_{k=0}^{N_2} g(k)b^k \cdot \sum_{m=0}^{N_1} f(m)b^m .$$

Tehát (8) és igy (7) is teljesül. A (6) egyenlőség (7)-ből a következő tételből közvetlenül következik.

2.3. **Tétel.** *Legyen*

$$\delta(n) = \begin{cases} 1, & \text{ha } n = 0 \\ 0, & \text{ha } n > 0 \end{cases} \tag{9}$$

és legyen a tetszőleges szám. Akkor bármely $f \in I$-re

$$(f * a\delta)(n) = af(n). \tag{10}$$

Bizonyitás.

$$(f*a\delta)(n) = \sum_{k=0}^{n} f(n-k)a\delta(k) = f(n)a\delta(0) = af(n).$$

Megjegyzés. A fenti tétel alapján (7)-ből (6) igy következik: A (9) alatt értelmezett δ függvény felhasználásával kapjuk minden $n \geqslant 0$ mellett, hogy

$$\left.\begin{aligned} &(S_b f)(n) = (S_b f)(0)\delta(n) \quad \text{és} \quad (S_b g)(n) = (S_b g)(0)\delta(n) \quad \text{és} \\ &(S_b(f*g))(n) = (S_b(f*g))(0)\delta(n) \end{aligned}\right\}$$

és igy (10) és (7) felhasználásával:

$$((S_b f) * (S_b g))(n) = (((S_b f)(0) \cdot \delta) * ((S_b f)(0) \cdot \delta))(n) =$$

$$= (S_b f)(0) \cdot (S_b g)(0)\delta(n) = (S_b(f*g))(0)\delta(0) = (S_b(f*g))(n).$$

Tehát valóban igaz a (6) egyenlőség.

Megjegyzés. Ha (10)-ben $a = 1$ akkor speciálisan kapjuk, hogy

$$f*\delta = f, \tag{11}$$

vagyis a δ függvény a konvolúciós szorzás egységeleme.

Tehát I egységelemes kommutatitv gyürü de nem test, mint arról könnyen meggyőződhetünk.

3. Az egészértékü aritmetikai függvények ekvivalenciája

3.1. **Definició.** I-ben bevezetünk egy ekvivalencia relációt, melyet a következőképpen értelmezünk: Legyen $f_1, f_2 \epsilon I$. f_1 ekvivalens f_2-vel, jelben: $f_1 \overset{b}{\sim} f_2$, akkor és csak akkor ha

$$S_b f_1 = S_b f_2 .$$

Megjegyzés. A 3.1. definicióból rögtön következik, hogy annak a szükséges és elégséges feltétele, hogy $f_1 \overset{b}{\sim} f_2$ legyen az, hogy

$$(1) \qquad \sum_{k=0}^{\mathrm{Supp} f_1}{}' f_1(k) b^k = \sum_{k=0}^{\mathrm{Supp} f_2}{}' f_2(k) b^k$$

teljesüljön.

Megjegyzés. A 3.1. definició alatt értelmezett ekvivalencia reláció természetesen b-től függ.
Ha két I-beli függvény egy adott b-re vonatkozóan ekvivalens, egy másik b-re vonatkozóan álatlában nem ekvivalens.
Ha például $f_1(0) = 1$, $f_1(1) = 1$, $\mathrm{Supp} f_1 = 1$ és $f_2(0) = 3$, $\mathrm{Supp} f_2 = 0$, akkor $f_1 \overset{2}{\sim} f_2$, de ha $b = 10$, akkor nem ekvivalensek, mert

$$\sum_{k=0}^{1}{}' f_1(k) 10^k = 11 \neq 3 = \sum_{k=0}^{0}{}' f_2(k) 10^k .$$

Ha a félreértés lehetősége kizárt, akkor az egyszerübb $\sim$ jelölést használjuk.

Megjegyzés. A 3.1. definicióban értelmezett reláció valóban ekvivalencia reláció, könnyen látható, hogy reflex, szimmetrikus, és tranzitiv, azaz

1.) $f \overset{b}{\sim} f$
2.) Ha $f_1 \overset{b}{\sim} f_2$ akkor $f_2 \overset{b}{\sim} f_1$
3.) Ha $f_1 \overset{b}{\sim} f_2$ és $f_2 \overset{b}{\sim} f_3$, akkor $f_1 \overset{b}{\sim} f_3$

A következőkben, mint eddig is, egy adott b-t mindig lerögzitettnek képzelünk, és igy a $\overset{b}{\sim}$ jelölés helyett a $\sim$ jelölést félreértés nélkül használhatjuk.

3.1. **Tétel.** *A* 3.1. *definicióban értelmezett ekvivalencia reláció kompatibilis az összeadásra és a konvolucios szorzásra nézve, vagyis, ha* $f_1 \sim f_2$ *és* $g_1 \sim g_2$,

$(f_1, f_2, g_1, g_2 \epsilon I)$, akkor

$$(2) \qquad f_1 + g_1 \sim f_2 + g_2$$

és

$$(3) \qquad f_1 * g_1 \sim f_2 * g_2 .$$

Bizonyitás. Mivel

$$(4) \qquad \sum_{k=0}^{\infty} f_1(k)b^k = \sum_{k=0}^{\infty} f_2(k)b^k$$

és

$$(5) \qquad \sum_{k=0}^{\infty} g_1(k)b^k = \sum_{k=0}^{\infty} g_2(k)b^k$$

ezért (4) és (5) alapján

$$\sum_{k=0}^{\infty} (f_1 + g_1)(k)b^k = \sum_{k=0}^{\infty} f_1(k)b^k + \sum_{k=0}^{\infty} g_1(k)b^k = \sum_{k=0}^{\infty} f_2(k)b^k + \sum_{k=0}^{\infty} g_2(k)b^k = \sum_{k=0}^{\infty} (f_2 + g_2)($$

azaz teljesül (2).

Most igazolni fogjuk, hogy (3) is fennáll. Mivel $f_1 \sim f_2$ és $g_1 \sim g_2$ ezért

$$(6) \qquad S_b f_1 = S_b f_2$$

és

$$(7) \qquad S_b g_1 = S_b g_2.$$

A 2.2 tétel felhasználásával (2-6 képlet) (6) és (7) miatt

$$S_b(f_1 * g_1) = (S_b f_1) * (S_b g_1) = (S_b f_2) * (S_b g_2) = S_b(f_2 * g_2),$$

vagyis (3) valóban igaz.

3.2. **Definició.** A 3.1. definició értelmében vett ekvivalencia meghatározza I-nek egy osztályozását, amennyiben egy osztályba soroljuk azokat az I-beli függvényeket, amelyek egymással ekvivalensek. Ilyen módon I felbomlik nem üres, páronként diszjunkt osztályok halmazára. Az osztályok halmazát I/b-vel jelöljük. Ha $f \in I$, akkor azt az osztályt melyben f benne van $\{f(n)\}_b$-vel jelöljük. Ha nem okoz félreértést, akkor $\{f(n)\}_b$ helyett rövidebben $\{f(n)\}$ -et is irhatunk.

3.3. **Definició.** Az I/b halmazban értelmezzük az összeadás és a szorzás müveletét a következőképpen:

Legyen $\{f(n)\}$ és $\{g(n)\}$ két tetszőleges I/b-be tartozó osztály, akkor az összeadást a

$$(8) \qquad \{f(n)\} + \{g(n)\} = \{f(n) + g(n)\}$$

és a szorzást a

$$(9) \qquad \{f(n)\}\{g(n)\} = \Big\{\sum_{k=0}^{n} f(n-k)g(k)\Big\} = \{(f * g)(n)\}$$

képlettel értelmezzük.

Megjegyzés. A (8) és (9) alatt értelmezett müveletek a 3.1. tétel következtében egyértelmüen vannak meghatározva.

3.2. **Tétel.** *Az I/b halmaz a* (8) *és* (9) *alatti müveletekre nézve kommutativ gyürüt alkot, mely az egész számok gyürüjével izomorf.*

Bizonyitás. Az hogy I/b kommutativ gyürü, az 1.2. tétel alapján könnyen látható. Azt kell tehát csak igazolni, hogy I/b leképezhető az egész számok E gyürüjére úgy, hogy a leképzés kölcsönösen egyértelmü és müvelettartó.

Megadjuk a leképzést: Ha $\{f(n)\} \in I/b$, akkor ehhez az osztályhoz az $(S_b f)(0) \in E$ egész számot rendeljük. Jelöljük az $\{f(n)\}$ osztályhoz rendelt számot $\overline{\{f(n)\}}$ -sal, akkor tehát azt irhatjuk:

$$(10) \qquad \overline{\{f(n)\}} = (S_b f)(0) = \sum_{k=0}^{\infty} f(k)b^k \in E.$$

A (10) leképzés egyértelmü, mert $\overline{\{f(n)\}}$ nem függ attól, hogy az $\{f(n)\}$ osztályból melyik f elemet választottuk ki.

Tegyük fel, hogy $\{f(n)\} = \{f_1(n)\}$. Akkor $f_1 \sim f$, vagyis a 3.1. megjegyzés alapján

$$\sum_{k=0}^{\infty} f_1(k)b^k = \sum_{k=0}^{\infty} f(k)b^k, \text{ tehát } \overline{\{f_1(n)\}} = \overline{\{f(n)\}}.$$

De a leképezés kölcsönösen egyértelmü is, mert ha valamilyen $\{f_1(n)\}$ és $\{f_2(n)\}$ osztályra $\overline{\{f_1(n)\}} = \overline{\{f_2(n)\}}$ teljesül, akkor (10) értelmezés szerint

$$\sum_{k=0}^{\infty} f_1(k)b^k = \sum_{k=0}^{\infty} f_2(k)b^k$$

teljesül, amiből a 3.1. megjegyzés alapján $f_1 \sim f_2$, azaz $\{f_1(n)\} = \{f_2(n)\}$ következik.

Megmutatjuk, hogy a (10) alatti leképezés müvelettartó, azaz

$$(11) \qquad \overline{\{f(n)\} + \{g(n)\}} = \overline{\{f(n)\}} + \overline{\{g(n)\}}$$

és

$$(12) \qquad \overline{\{f(n)\}\{g(n)\}} = \overline{\{f(n)\}} \cdot \overline{\{g(n)\}}.$$

Valóban, a 2.1. tétel és (8) és (10) felhasználásával kapjuk, hogy

$$\overline{\{f(n)\}} + \overline{\{g(n)\}} = \overline{\{f(n) + g(n)\}} = (S_b(f + g))(0) =$$
$$= (S_b f)(0) + (S_b g)(0) = \overline{\{f(n)\}} + \overline{\{g(n)\}},$$

tehát (11) igaz.

(12) igazolásához a 2.2 tétel (7) képletét, valamint (9)-et és (10)-et használjuk fel:

$$\overline{\{f(n)\}\{g(n)\}} = \overline{\{(f * g)(n)\}} = (S_b(f * g))(0) = (S_b f)(0) \cdot (S_b g)(0) =$$
$$= \overline{\{f(n)\}} \cdot \overline{\{g(n)\}},$$

tehát (12) is igaz. Ezzel megmutattuk, hogy I/b és E izomorf.

4. Az átvitelt végrehajtó Tb transzformáció

4.1. **Definició.** Aritmetikai függvénynek nevezzük az f függvényt, ha f értelmezési tartománya a nem negativ egész számok halmaza és $f(n)$ valós számértéket vesz fel minden n mellett. Az aritmetikai függvények halmazát A-val jelöljük.

4.2. **Definició.** Az A halmazon értelmezünk egy T_b transzformációt, mely minden $f \epsilon A$ függvényhez egy $T_b f \epsilon A$ függvényt rendel. A T_b transzformációt a következő egyenletrendszer segitségével értelmezzük:

$$\text{(1)} \quad \begin{aligned} &\text{a)} \quad s(0) = f(0) \\ &\text{b)} \quad s(n+1) = f(n+1) + \left[\frac{s(n)}{b}\right] \\ &\text{c)} \quad (T_b f)(n) = s(n) - b\left[\frac{s(n)}{b}\right], \end{aligned}$$

ahol $s(n)$ egy egész értékeket felvevő segédfüggvény, melyet az f ismeretében az (1.a.) ill. (1.b.) egyenletek egyértelmüen meghatároznak. s-et az f-hez tartozó állapotfüggvénynek nevezzük.

4.3. **Definició.** Jelöljük I_+-szal azoknak a függvényeknek a halmazát, melyek értelmezési tartománya a nemnegativ egészek halmaza, és ezen egész számértékeket vesznek fel és a következő tulajdonságoknak tesznek eleget:

I. Minden $f \epsilon I_+$ függvényre $0 \leqslant f(n) < b \quad (n = 0,1,2, \ldots)$

II. Minden $f \epsilon I_+$ függvényhez létezik Supp$f$, ha $f \neq 0$, vagyis minden $f \neq 0$, $f \epsilon I$-hez – f-től függően – létezik N nemnegativ egész, hogy $f(N) \neq 0$, de minden $n > N$ egész számra $f(n) = 0$.

III. Ha $f = 0$, akkor $f \epsilon I_+$.

Megjegyzés. Világos, hogy $I_+ \subset I$.

4.4. **Definició.** Jelöljük I_--szal azoknak a nemnegativ egész számok halmazán értelmezett, egész értékeket felvevő függvényeknek a halmazát, amelyek a következő tulajdonságoknak tesznek eleget:

(*i*) Minden $g \in I_-$-függvényre $0 \leqslant g(n) \leqslant b$.

(*ii*) Minden $g \in I_-$-hoz – g-től függően – létezik olyan M nemnegativ egész szám, hogy ha $n \geqslant M$, akkor $g(n) = b - 1$.

Megjegyzés. Világos, hogy $I_- \cap I = \emptyset$, vagyis az I_--ba tartozó függvények nem I-beliek és forditva, ha $f \in I$ akkor $f \notin I_-$.

4.1. **Tétel.** Legyen $f \in I$. Ha $(S_b f)(0) \geqslant 0$, akkor $T_b f \in I_+$ ha $(S_b f)(0) < 0$, akkor $T_b f \in I_-$.

Bizonyitás. Először megmutatjuk, hogy ha $f \in I$, akkor minden n-re

$$0 \leqslant (T_b f)(n) < b \,. \tag{2}$$

Valóban, mivel minden valós x-re $0 \leqslant x - [x] < 1$, ezért az (1.c.) egyenletből adódó

$$\frac{(T_b f)(n)}{b} = \frac{s(n)}{b} - \left[\frac{s(n)}{b}\right]$$

összefüggésből $0 \leqslant \dfrac{(T_b f)(n)}{b} < 1$ adódik, ahonnan (2) következik.

Ezzel megmutattuk, hogy $T_b f$ eleget tesz a 4.3. definició I. és a 4.4. definició (*i*) tulajdonságának minden n-re. Azt könnyü látni (1)-ből hogy $(T_b f)(n)$ egész szám minden n-re.

Most megmutatjuk, hogy ha $(S_b f)(0) \geqslant 0$, akkor $T_b f$-re 4.3. definició I. tulajdonsága mellett teljesül a 4.3. definició II. tulajdonsága is, és igy $T_b f \in I_+$, ha pedig $(S_b f)(0) < 0$, akkor $T_b f$-re teljesül a 4.4. definició (*i*) tulajdonsága mellett a 4.4. definició (*ii*) tulajdonsága is, és igy $T_b f \in I_-$.

Ehhez szükségünk van a következő lemmára:

4.1. **Lemma.** *Legyen* $f \in I$ *és legyen* $s = s(n)$ *az* (1.b.) *egyenletnek az* (1.a.) *kezdeti feltétel melletti megoldása. Ha* $s(\mathrm{Supp} f) \geqslant 0$, *akkor létezik olyan* N_1, *hogy ha* $n \geqslant N_1$, *akkor* $s(n) = 0$, *ha pedig* $s(\mathrm{Supp} f) < 0$, *akkor létezik olyan* N_2 *egész szám, hogy ha* $n \geqslant N_2$, *akkor* $s(n) = -1$.

Bizonyitás. Legyen $N_0 = \mathrm{Supp} f$. Akkor $f(N_0 + 1) = 0$ miatt (1.b)- ből azt kapjuk, hogy

$$s(N_0 + 1) = \left[\frac{s(N_0)}{b}\right] \leqslant \frac{s(N_0)}{b} \,. \tag{3}$$

Megmutatjuk teljes indukcióval, hogy tetszőleges k természetes számra

$$s(N_0 + k) \leqslant \frac{s(N_0)}{b^k} \,. \tag{4}$$

$k = 1$-re (4) a már igazolt (3) egyenlőtlenségbe megy át. Tegyük fel, hogy (4) teljesül $k \geqslant 1$-re. Akkor $f(N_0 + k + 1) = 0$, (1.b.) és (4) felhasználásával kapjuk:

$$s(N_0 + k + 1) = \left[\frac{s(N_0 + k)}{b}\right] \leqslant \frac{s(N_0 + k)}{b} \leqslant \frac{s(N_0)}{b^{k+1}},$$

tehát igaz (4). Ha $s(N_0) \geqslant 0$, akkor teljes indukcióval (1.b.) alapján beláthatjuk, hogy $s(N_0 + k) \geqslant 0$ minden k természetes számra. Ha K olyan nagy, hogy $\frac{s(N_0)}{b^K} < 1$, akkor (4)-ből $0 \leqslant s(N_0 + K) < 1$ következik, és mivel $s(N_0 + K)$ egész szám, ez csak úgy teljesülhet, hogy $s(N_0 + K) = 0$. Legyen $N_1 = N_0 + K$. Akkor $s(N_1) = 0$ és igy minden $k \geqslant 0$ egész szám (1.b.)-ből $s(N_1 + k) = 0$ következik, vagyis, ha

$$n \geqslant N_1, \text{ akkor } s(n) = 0.$$

Legyen most $s(N_0) < 0$. Akkor minden k természetes számra

$$(5) \qquad \frac{s(N_0)}{b^k} - s(N_0 + k) < \frac{b}{b-1}.$$

$k = 1$-re ugyanis (5) az $\frac{s(N_0)}{b} - s(N_0 + 1) < \frac{b}{b-1}$ egyenlőtlenségbe megy át, ami $s(N_0 + 1) = \left[\frac{s(N_0)}{b}\right]$ miatt nyilván igaz. Tegyük fel, hogy (5) igaz $k \geqslant 1$-re, akkor (5)-ből kapjuk:

$$(6) \qquad \frac{s(N_0)}{b^{k+1}} - \frac{s(N_0 + k)}{b} < \frac{1}{b}\,\frac{b}{b-1} = \frac{1}{b-1}$$

Mivel $\left[\frac{s(N_0 + k)}{b}\right] = s(N_0 + k + 1)$, ezért

$$(7) \qquad \frac{s(N_0 + k)}{b} - s(N_0 + k + 1) < 1.$$

A (6) és (7) egyenlőtlenségek megfelelő oldalainak az összeadása után kapjuk, hogy

$$\frac{s(N_0)}{b^{k+1}} - s(N_0 + k + 1) < 1 + \frac{1}{b-1} = \frac{b}{b-1}.$$

Tehát (5) igaz minden k-ra. (5)-ből azt kapjuk, hogy

$$\frac{s(N_0)}{b^k} - s(N_0 + k) < \frac{b}{b-1} = 1 + \frac{1}{b-1}, \text{ és innen}$$

$$(8) \qquad s(N_0 + k) > -1 - \frac{1}{b-1} + \frac{s(N_0)}{b^k} \geqslant -2 + \frac{s(N_0)}{b^k} .$$

Ha most K olyan nagy, hogy $\frac{s(N_0)}{b^K} > -1$, akkor

$$s(N_0 + K) > -2 + \frac{s(N_0)}{b^K} > -3,$$

ahonnan

$$\frac{s(N_0 + K)}{b} > -\frac{3}{b} \geqslant -\frac{3}{2},$$

tehát

$$(9) \qquad s(N_0 + K + 1) = \left[\frac{s(N_0 + K)}{b}\right] \geqslant -2.$$

(9)-ből (1.b.) alapján felhasználva, hogy $s(n) < 0$, ha $s(N_0) < 0$, és $n \geqslant N_0$, továbbá azt, hogy

$$\frac{s(N_0 + K + 1)}{b} \geqslant -\frac{2}{b} \geqslant -1,$$

kapjuk, hogy

$$(10) \qquad 0 > s(N_0 + K + 2) = \left[\frac{s(N_0 + K + 1)}{b}\right] = -1.$$

Legyen $N_2 = N_0 + K + 2$. Akkor $s(n+1) = \left[\frac{s(n)}{b}\right]$, $(n \geqslant N_2)$ egyenlőségből teljes indukcióra kapjuk, hogy $s(n) = -1$. Ugyanis az állitás $n = N_2$-re igaz. Tegyük fel, hogy $n \geqslant N_2$ mellett $s(n) = -1$. Igy

$$\left[\frac{s(n)}{b}\right] = \left[\frac{-1}{b}\right] = -1.$$

Tehát, ha $n \geqslant N_2$, akkor $s(n) = -1$.

Ezzel a lemmát igazoltuk.

A 4.1. tétel bizonyitásának a folytatása:
Az (1) egyenletek felhasználásával kapjuk, hogy

$$\sum_{n=0}^{N} (T_b f)(n) b^n = \sum_{n=0}^{N} s(n) b^n - \sum_{n=0}^{N} b^{n+1} \left[\frac{s(n)}{b} \right] =$$

$$= f(0) + \sum_{n=1}^{N} f(n) b^n + \sum_{n=1}^{N} b^n \left[\frac{s(n-1)}{b} \right] - \sum_{n=0}^{N} b^{n+1} \left[\frac{s(n)}{b} \right] =$$

$$= \sum_{n=0}^{N} f(n) b^n + \sum_{n=0}^{N-1} b^{n+1} \left[\frac{s(n)}{b} \right] - \sum_{n=0}^{N-1} b^{n+1} \left[\frac{s(n)}{b} \right] - b^{N+1} \left[\frac{s(N)}{b} \right]$$

vagyis

$$(11) \qquad \sum_{n=0}^{N} (T_b f)(n) b^n = \sum_{n=0}^{N} f(n) b^n - b^{N+1} \left[\frac{s(N)}{b} \right].$$

Tegyük fel, hogy $(S_b f)(0) = \sum_{k=0}^{Supp\, f} f(k) b^k \geqslant 0$, és legyen a rövidség kedvéért $Supp f = N_0$. Akkor $s(N_0) \geqslant 0$. Ha ugyanis $s(N_0) < 0$, akkor van olyan N_2, hogy ha $n \geqslant N_2$, akkor $s(n) = -1$. Legyen $N > \max(N_0, N_2)$, akkor (11) átmegy a

$$(12) \qquad \sum_{n=0}^{N} (T_b f)(n) b^n = \sum_{n=0}^{N} f(n) b^n + b^{N+1} = (S_b f)(0) + b^{N+1}$$

egyenlőségbe. Mivel (2) miatt $(T_b f)(n) \leqslant b - 1$, (12)-ből a következő egyenlőtlenséget kapjuk

$$(13) \qquad (S_b f)(0) + b^{N+1} \leqslant \sum_{n=0}^{N} (b-1) b^n = b^{N+1} - 1,$$

ahonnan az $(S_b f)(0) \leqslant -1$ ellentmondásra jutunk.
Tehát, ha $(S_b f)(0) \geqslant 0$, akkor $s(N_0) \geqslant 0$. De akkor van a 4.1. lemma szerint olyan N_1, hogy ha $n \geqslant N_1$, akkor $s(n) = 0$. De akkor (1.c.) alapján azt kapjuk, hogy $n \geqslant N_1$ esetén

$$(T_b f)(n) = 0$$

Ezzel megmutattuk, hogy $T_b f \in I_+$.

Tegyük most, fel hogy $(S_b f)(0) < 0$. Akkor $s(Supp f) < 0$. Ha ugyanis $s(Supp f) \geqslant 0$ volna, akkor a 4.1. lemma szerint létezne N_1, hogy $s(N) = 0$, ha $N \geqslant N_1$. Legyen $N \geqslant \max(N_1, N_0)$, akkor igy (11) átmenne a

$$(14) \qquad \sum_{n=0}^{N} (T_b f)(n) b^n = \sum_{n=0}^{N} f(n) b^n = (S_b f)(0)$$

egyenlőségbe. Azonban (2) szerint $(T_b f)(n) \geqslant 0$ és igy (14)-ből az $(S_b f)(0) \geqslant 0$ ellentmondásra jutnánk. Tehát $s(Supp f) < 0$. De akkor a 4.1. lemma alapján létezik N_2, hogy minden $n \geqslant N_2$ mellett $s(n) = -1$. Igy tehát $n \geqslant N_2$ esetén (1c)-ből kapjuk, hogy

$$(T_b f)(n) = s(n) - b\left[\frac{s(n)}{b}\right] = -1 + b = b - 1.$$

Tehát $(S_b f)(0) < 0$ esetén $T_b f \epsilon I_-$.

4.2. **Tétel.** *Legyen* $f, g \epsilon I$. *Akkor* $f \overset{b}{\sim} g$ *akkor és csak akkor teljesül, ha*

$$T_b f = T_b g \tag{15}$$

Bizonyítás. A 3.1. megjegyzés alapján azt kell igazolni, hogy ha $f,g \epsilon I$, akkor

$$(S_b f)(0) = (S_b g)(0) \tag{16}$$

akkor és csak akkor igaz, ha (15) igaz.

Tegyük fel először, hogy (16) teljesül. Legyen s_1, az f-hez, s_2 a g-hez tartozó állapotfüggvény. A 4.1. lemma értelmében létezik olyan N_1^* és N_2^*, hogy $n > N_1^*$ esetén $s_1(n) = 0$ vagy $s_1(n) = -1$ aszerint, hogy $(S_b f)(0) \geqslant 0$ vagy $(S_b f)(0) < 0$.

Ugyanigy, g-hez is létezik egy N_2^* úgy, hogy $s_2(n) = 0$, ha $(S_b g)(0) \geqslant 0$ és $n \geqslant N_2^*$, vagy ha $(S_b g)(0) \leqslant 0$, akkor $s_2(n) = -1$ minden $n \geqslant N_2^*$ mellett. Legyen $N^* = \max(N_1^*, N_2^*)$, akkor (16) miatt $N \geqslant N^*$ esetén vagy $s_1(N) = s_2(N) = 0$ vagy $s_1(N) = s_2(N) = -1$.

Tehát minden esetben

$$\left[\frac{s_1(N)}{b}\right] = \left[\frac{s_2(N)}{b}\right]. \tag{17}$$

Legyen $N \geqslant \max(N^*, \mathrm{Supp} f, \mathrm{Supp} g)$ és ilyen N mellett tekintsünk a (11) egyenlőséget f-re és g-re:

$$\sum_{n=0}^{N} (T_b f)(n) b^n = (S_b f)(0) - b^{N+1}\left[\frac{s_1(N)}{b}\right] \tag{18}$$

$$\sum_{n=0}^{N} (T_b g)(n) b^n = (S_b g)(0) - b^{N+1}\left[\frac{s_2(N)}{b}\right]. \tag{19}$$

A (18)-ból és (19)-ből (16)-ra és (17)-re való tekintettel kapjuk, hogy

$$\sum_{n=0}^{N} (T_b f)(n) b^n = \sum_{n=0}^{N} (T_b g)(n) b^n. \tag{20}$$

Mivel a 4.1. tétel miatt $T_b f, T_b g \epsilon I_+ \cup I_-$, a (20) baloldala is meg a jobboldala is ugyanannak a számnak a b alapú számrendszerben felírt alakja. Ez a felírás egyértelmű, ezért (20)-ból

$$(T_b f)(n) = (T_b g)(n) \tag{21}$$

következik minden $0 \leqslant n \leqslant N$ számra. Mivel N tetszőleges nagy lehet, (21) igaz minden n-re,

és igy igaz (15).

Tegyük most fel, hogy (15) igaz, vagyis (21) teljesül minden n-re. Akkor (1.c.) és (21) alapján

$$(22) \qquad s_1(n) - b\left[\frac{s_1(n)}{b}\right] = s_2(n) - b\left[\frac{s_2(n)}{b}\right] \qquad (n = 0,1,2,\ldots)$$

ahol s_1, az f-hez, s_2 a g-hez tartozó állapotfüggvény, (22)-ből kapjuk:

$$(23) \qquad \frac{s_1(n) - s_2(n)}{b} = \left[\frac{s_1(n)}{b}\right] - \left[\frac{s_2(n)}{b}\right] \qquad (n = 0,1,2,\ldots)$$

Legyen $N > \max(N_1^*, N_2^*, \mathrm{Supp}f, \mathrm{Supp}g)$. Akkor ilyen N mellett tekintve a (18) és (19) egyenlőségeket (21) figyelembevételével kapjuk:

$$(S_b f)(0) - b^{N+1}\left[\frac{s_1(N)}{b}\right] = (S_b g)(0) - b^{N+1}\left[\frac{s_2(N)}{b}\right],$$

ahonnan átrendezéssel adódik

$$(S_b f)(0) - (S_b g)(0) = b^{N+1}\left(\left[\frac{s_1(N)}{b}\right] - \left[\frac{s_2(N)}{b}\right]\right)$$

és innen (23)-ra való tekintettel

$$(24) \qquad (S_b f)(0) - (S_b g)(0) = b^N(s_1(N) - s_2(N)).$$

Mivel (24) bal oldala nem függ a N számtól, ezért a jobboldal sem függ N-től. De mivel a N-re tett feltevés miatt $s_1(N) - s_2(N)$ sem függ N-től, (24) jobboldala csak akkor lehet konstans, ha $s_1(N) - s_2(N) = 0$. De akkor viszont (24)-ből (16) következik, amit bizonyitani kellett.

4.3. **Tétel.** *Minden* $n \geqslant 0$ *egészre és minden* $f \in A$ *függvényre igaz*

$$(25) \qquad (T_b^2 f)(n) = (T_b f)(n), \qquad \textit{azaz } T_b \textit{ idempotens operátor :}$$

$$(26) \qquad T_b^2 = T_b.$$

Bizonyitás. Legyen $f \in A$, akkor a $g(n) = (T_b^2 f)(n) = (T_b(T_b f))(n)$ függvény értékeit az (1) egyenletrendszerből kaphatjuk meg:

$$(27) \qquad s(0) = (T_b f)(0)$$

$$(28) \qquad s(n+1) = (T_b f)(n+1) + \left[\frac{s(n)}{b}\right]$$

$$(29) \qquad (T_b^2 f)(n) = s(n) - b\left[\frac{s(n)}{b}\right]$$

Teljes indukcióval bebizonyitjuk, hogy

(30) $\quad s(n) = (T_b f)(n) \quad (n = 0,1,2,\ldots)$

$n = 0$-ra (27) miatt igaz az álllitás.

Tegyük fel, hogy (30) igaz $n \geqslant 0$ mellett. Mivel (2) szerint $0 \leqslant (T_b f)(n) < b$, igy $\left[\frac{(T_b f)(n)}{b}\right] = 0$

Tehát (28)-ból (30) felhasználásával kapjuk, hogy

$$s(n+1) = (T_b f)(n+1) + \left[\frac{s(n)}{b}\right] = (T_b f)(n+1) + \left[\frac{(T_b f)(n)}{b}\right] = (T_b f)(n+1).$$

Ezzel igazoltuk a (30)-egyenlőséget. Akkor viszont (29) alapján

$$(T_b^2 f)(n) = (T_b f)(n) - b\left[\frac{(T_b f)(n)}{b}\right] = (T_b f)(n).$$

Ezzel igazoltuk a (25) egyenlőséget és igy a (26)-ot is.

4.4.**Tétel.** *Ha* $f,g \in I$, *akkor*

(31) $\quad T_b(f+g) = T_b(T_b f + T_b g)$

és

(32) $\quad T_b(f * g) = T_b(T_b f * T_b g).$

Bizonyitás. Elöszőr megmutatjuk, hogy ha $f \in I$, akkor $f \overset{b}{\sim} T_b f$.

A 4.3. tétel alapján $T_b f = T_b^2 f = T_b(T_b f)$ és innen a 4.2 tétel alkalmazásával kapjuk, hogy

(33) $\quad f \overset{b}{\sim} T_b f$

Mivel $g \in I$, ugyanigy fennáll

(34) $\quad g \overset{b}{\sim} T_b g$

is. Igy a 3.1. tétel alkalmazásával kapjuk, hogy

(35) $\quad f + g \overset{b}{\sim} T_b f + T_b g$

és

(36) $\quad f * g \overset{b}{\sim} T_b f * T_b g$

(35) és (36)-ból a 4.2. tétel alkalmazásával kapjuk (31) és (32)-t.

4.5.**Tétel.** *Ha* $f \in I_+ \cup I_-$, *akkor*

(37) $\quad T_b f = f.$

Bizonyitás. Legyen $f \epsilon I_+ \cup I_-$. Akkor minden $n \geqslant 0$ egész számra

(38) $\quad 0 \leqslant f(n) < b.$

Tehát ha s az f-hez tartozó állapotfüggvény, akkor

$$\left[\frac{s(0)}{b}\right] = \left[\frac{f(0)}{b}\right] = 0.$$

Megmutatjuk

(39) $\quad \left[\frac{s(n)}{b}\right] = 0.$

$n = 0$-ra a (39) egyenlőséget már beláttuk. Tegyük fel, hogy (38) igaz $n \geqslant 0$ esetén. Akkor (1.b.) alapján

$$s(n + 1) = f(n + 1)$$

tehát (38) miatt $\left[\frac{s(n+1)}{b}\right] = \left[\frac{f(n+1)}{b}\right] = 0$. Tehát (39) igaz minden n-re. De akkor (1.b.) és (1.c.) alapján minden n-re

$$(T_b f)(n) = s(n) = f(n)$$

vagyis (37) valóban fennáll.

5. Algoritmus hosszú egész számnak rövid számmal való osztására

5.1. **Tétel.** *Legyen* $f \epsilon I_+$ *és valamilyen egész szám legyen* $a \neq 0$. *Akkor az*

$$ax(0) - bt(0) = f(0)$$

(1)

$$ax(n) - bt(n) = f(n) - t(n - 1)$$

egyenletrendszernek a

(2) $\quad 0 \leqslant x(n) < b \qquad (x(n), t(n) \epsilon E^*) \qquad (n = 0,1,2, \ldots)$

feltételek mellett legfeljebb egyetlen $x(n)$ *és* $t(n)$ *függvény megoldáspárja létezik.*

Bizonyitás. Tegyük fel, hogy az $x(n)$ és $t(n)$ függvénypár mellett az $x'(n), t'(n)$ függvénypár is megoldása (1)-nek ugy, hogy

(3) $\quad 0 \leqslant x'(n) < b \qquad (n = 0,1,2, \ldots)$

teljesül. Mivel $x'(n)$ és $t'(n)$ is megoldása (1)-nek, ezért

$$(4)\qquad \begin{cases} ax'(0) - bt'(0) = f(0) \\ \\ ax'(n) - bt'(n) = f(n) - t'(n-1) \end{cases} \qquad (n = 1,2,\ldots)$$

Legyen $y(n) = x(n) - x'(n)$, $z(n) = t(n) - t'(n)$ $(n = 0,1,2,\ldots)$. Akkor az (1) és (4) egyenletekből kapjuk, hogy

$$(5)\qquad \begin{cases} ay(0) - bz(0) = 0 \\ \\ ay(n) - bz(n) = -z(n-1) \end{cases} \qquad (n = 1,2,\ldots)$$

Mivel az $ay - bz = 0$ diofantikus egyenlet álatlános megoldása $y = kb$ és $z = ka$, ahol $k = 0, \pm 1, \pm 2, \ldots$ ezért az (5) egyenletrendszer első egyenletéből $y(0) = kb$ és $z(0) = ka$ alaku lehet.

A (2) és (3) feltételből következik, hogy

$$(6)\qquad |y(n)| < b \qquad (n = 0,1,2,\ldots)$$

Tehát speciálisan $n = 0$ mellett

$$|k|b = |kb| = |y(0)| < b$$

ahonnan azt kapjuk, hogy $|k| < 1$. Ez csak $k = 0$ mellett lehetséges, és igy $y(0) = 0$, és $z(0) = 0$. Megmutatjuk, hogy

$$(7)\qquad y(n) = 0$$

és

$$(8)\qquad z(n) = 0$$

teljesül minden $n = 0,1,2,\ldots$ mellett. $n = 0$ esetre beláttuk az állitást. Tegyük fel, hogy valamilyen $n \geqslant 1$ mellett $y(n-1) = 0$ és $z(n-1) = 0$. Az indukciós bizonyitáshoz azt kell megmutatni, hogy akkor (7) és (8) is teljesül. Valóban, az indukciós feltevés miatt az (5) alatti második egyenlet ilyen alaku lesz;

$$(9)\qquad ay(n) - bz(n) = 0$$

Tehát $y(n) = kb$ és $z(n) = ka$ és (6) miatt $k = 0$, vagyis valóban igaz (7) és (8). Igy tehát azt kapjuk, hogy $x(n) = x'(n)$ és $t(n) = t'(n)$ minden n-re. Ezzel a tételt igazoltuk.

Megjegyzés. Jól ismert, hogy az

$$ax - bt = c$$

diofanatikus egyenletnek akkor és csak akkor van megoldása, ha a és b legnagyobb közös osztója c-nek is osztója. Igy tehát az (1) egyenletrendszernek biztosan nincs megoldása, ha (a,b) nem osztója $f(c)$-nak. Az (1) egyenletrendszernek mindig van megoldása, ha a és b relativ prim, mert akkor $(a,b) = 1$ minden szám osztója.

5.2. **Tétel.** *Legyen* $f \epsilon I_+$ *és* $M = \sum_{k=0}^{\mathrm{Supp} f} f(k)b^k \geqslant 0$.

Ha $a > 0$ osztója M-nek, akkor létezik az 5.1. tétel szerint

$$ax(0) - bt(0) = f(0)$$

$$(n = 1,2, \ldots)$$

$$ax(n) - bt(n) = f(n) - t(n-1)$$

egyenletrendszernek olyan $x(n)$, $t(n)$ megoldáspárja, amelyre a

$$0 \leqslant x(n) < b$$

feltétel teljesül. Ekkor $x \epsilon I_+$ és

$$(10) \qquad \sum_{k=0}^{\mathrm{Supp} x} x(k)b^k = \frac{M}{a}$$

Bizonyitás. Legyen $m = \frac{M}{a}$ és $\mu \epsilon I_+$ olyan, hogy

$$(11) \qquad m = \sum_{k=0}^{\mathrm{Supp}\mu} \mu(k)b^k$$

vagyis (11) az m szám b alapu számrendszerben felirt alakja. De akkor $a > 0$ miatt

$$S_b(a\mu)(0) = \sum_{k=0}^{\mathrm{Supp}\mu} a\mu(k)b^k = am = M = \sum_{k=0}^{\mathrm{Supp} f} f(k)b^k = (S_b f)(0),$$

tehát $f \overset{b}{\sim} a\mu$. Igy a 4.2. tétel alkalmazásával kapjuk, hogy

$$(13) \qquad T_b f = T_b(a\mu).$$

$f \epsilon I_+$ miatt a 4.5. tétel alapján $T_b f = f$ és igy

$$(14) \qquad f = T_b(a\mu)$$

Legyen s az $a\mu$ függvényhez tartozó állapotfüggvény. Megmutatjuk, hogy akkor $x(n) = \mu(n)$ és $t(n) = \left[\frac{s(n)}{b}\right]$ kielégiti az (1) egyenletrendszert, és a (2) egyenlőtlenség is teljesül. Valóban a 4.2. definició szerint

(15) $s(0) = a\mu(0)$

(16) $s(n+1) = a\mu(n+1) + \left[\frac{s(n)}{b}\right]$

(17) $T_b(a\mu)(n) = s(n) - b\left[\frac{s(n)}{b}\right]$

Igy tehát (14)-re való tekintettel kapjuk, hogy

$$f(0) = T_b(a\mu)(0) = s(0) - b\left[\frac{s(0)}{b}\right] = a\mu(0) - b\left[\frac{s(0)}{b}\right] = ax(0) - bt(0),$$

vagyis, az (1) első egyenlete teljesül. Legyen $n \geqslant 1$. Akkor (14), (16) és (17) miatt

$$f(n) = T_b(a\mu)(n) = s(n) - b\left[\frac{s(n)}{b}\right] = a\mu(n) + \left[\frac{s(n-1)}{b}\right] - b\left[\frac{s(n)}{b}\right] =$$
$$= ax(n) + t(n-1) - bt(n),$$

ahonnan átrendezéssel kapjuk az (1) egyenleteit $n = 1,2,\ldots$ esetére. Az, hogy a (2) egyenlőtlenség $x = \mu$ esetén fennáll, az $a\mu \epsilon I_+$ feltétel miatt nyilván teljesül. Az is nyilvánvaló, hogy $x = \mu$ mellett (10) is fennáll (a (11) egyenlőség és $m = \frac{M}{a}$ miatt). Ezzel a tételt igazoltuk.

6. Algoritmus hosszú egész számnak hosszú egész számmal való osztására

A következőkben adott f és g mellett vizsgáljuk a

(1) $(g * x)(n) - bt(n) = f(n) - t(n-1) \qquad (t(-1) = 0;\ n = 0,1,\ldots)$

egyenleteknek a

(2) $0 \leqslant x(n) < b \qquad (x(n),\ t(n)$ egész; $n = 0,1,2,\ldots)$

feltételek melletti megoldásait. Az (1) egyenletrendszer az (5.1.) egyenletrendszer általánositása. Ugyanis, ha speciális $g(0) = a$ és $Suppg = 0$ akkor (1) átmegy az (5.1) egyenletrendszerbe.

6.1. **Tétel.** *Legyenek f és g egész értékeket felvevő függvények. Ha $g \neq 0$, akkor legfeljebb egy olyan $x(n)$, $t(n)$ függvénypár létezik, amelyre az* (1) *egyenletek fennállnak és ugyanakkor a* (2) *feltételek is teljesülnek.*

Bizonyitás. Tegyük fel, hogy $x(n)$-re és $t(n)$-re teljesül (1) és (2) és ugyanakkor $x'(n)$ és $t'(n)$-re is fennáll, hogy

(3) $(g * x')(n) - bt'(n) = f(n) - t'(n-1) \qquad (t'(-1) = 0;\ n = 0,1,\ldots)$

és

(4) $0 \leqslant x'(n) < b \qquad (x'(n),\ t'(n)$ egész; $n = 0,1,\ldots)$

Legyen $y = x - x'$ és $z = t - t'$. Akkor (1)-ből és (3)-ból kapjuk, hogy

(5) $\qquad (g * y)(n) - bz(n) = - z(n - 1) \qquad (z(- 1) = 0; n = 0,1, \ldots)$

és (2)-ből (4)-ből kapjuk:

(6) $\qquad |y(n)| < b \qquad (n = 0,1, \ldots)$

Az (5) egyenletet részletesen kiirva kapjuk, hogy

(7) $\qquad \sum_{k=0}^{n} g(n - k)y(k) - bz(n) = - z(n - 1) \qquad (z(- 1) = 0; n = 0,1, \ldots)$

$n = 0$ esetén a (7) egyenlet a

(8) $\qquad g(0)y(0) - bz(0) = 0$

egyneletbe megy át. Megmutatjuk, hogy $y(0) = z(0) = 0$. Ez $g(0) \neq 0$ esetén (8)-ból (6) miatt következik. Ha $g(0) = 0$, akkor (8)-ból egyenlőre csak az következik, hogy $z(0) = 0$. Tegyük fel, hogy $g(0) = 0$. Mivel $g \not\equiv 0$, létezik olyan $r > 0$, hogy $g(r) \neq 0$, de ha $0 \leqslant n < r$, akkor $g(n) = 0$. Ha tehát $0 \leqslant n < r$, akkor a (7) egyenlet a

(9) $\qquad - bz(n) = - z(n - 1) \qquad (n = 0, \ldots, r - 1)$

egyenletbe megy át, ahonnan (teljes indukcióval) következik, hogy

(10) $\qquad z(n) = 0, \qquad$ ha $\qquad 0 \leqslant n < r.$

Helyettesitsük (7)-be n helyébe r-et, akkor (10)-re való tekintettel azt kapjuk, hogy

(11) $\qquad g(r)y(0) - bz(r) = - z(r - 1) = 0.$

Innen $g(r) \neq 0$ és (6) miatt az következik, hogy

(12) $\qquad y(0) = z(r) = 0$

Azt akarjuk teljes indukcióval megmutatni, hogy $y(n) = z(n + r) = 0$ minden n-re igaz. Ezt $n = 0$ esetére már beláttuk.

Tegyük fel, hogy valamilyen $n > 0$ esetén

(13) $\qquad y(0) = y(1) = \ldots = y(n - 1) = 0, \qquad z(0) = z(1) = \ldots = z(n + r - 1) = 0$

igaz. Helyettesitsünk (7)-be n helyébe $n + r$-et akkor

(14) $\qquad g(n + r - k)y(k) = 0,$

ha $k < n$, mert akkor $y(k) = 0$. Ha pedig $n < k \leqslant n + r$, akkor $0 \leqslant n + r - k < r$ és igy $g(n + r - k) = 0$, aminek következtében (14) ismét fennáll. Igy tehát a (7) egyenlet a következő egyenletre redukálódik:

(15) $\quad g(r)y(n) - bz(n + r) = - z(n + r) = - z(n + r - 1)$

Az indukciós feltevés miatt $z(n + r - 1) = 0$, és igy a (15) egyenletből $g(r) \neq 0$ és (6) miatt $y(n) = 0$ és $z(n + r) = 0$ következik. Ezek megmutaták, hogy minden n-re

$$x(n) - x'(n) = y(n) = 0 \quad \text{és és} \quad t(n) - t'(n) = 0,$$

vagyis igazoltuk a tételt.

Megjegyzés. A 6.1 tétel egyik következménye: Adott f és g mellett, ha $g \neq 0$, akkor legfeljebb egyetlen olyan $x \in I_+ \cup I_-$ létezik, hogy valamely $t = t(n)$ mellett az (1) egyenletek teljesüljenek.

6.2 **Tétel.** *Legyen* $f,g \in I$, $g \neq 0$. *Ha van olyan* $x \in I$, *amely megoldása* (1)*-nek es teljesülnek a* (2) *feltételek is, akkor*

$$\sum_{k=0}^{\mathrm{Supp}f} f(k)b^k \quad \textit{osztható az} \quad \sum_{k=0}^{\mathrm{Supp}g} g(k)b^k \quad \textit{számmal, és}$$

(16) $$\sum_{k=0}^{\mathrm{Supp}x} x(k)b^k = \frac{\sum_{k=0}^{\mathrm{Supp}f} f(k)b^k}{\sum_{k=0}^{\mathrm{Supp}g} g(k)b^k}$$

Bizonyitás. Legyen x az (1) egyenlet megoldása, és legyen

(17) $\quad N > \max(\mathrm{Supp}f, \mathrm{Supp}\ (g * x))$.

Az (1) egyenlet mindkét oldalát megszorozva b^n-vel és összegezve $n = 0$-tól $n = N$-ig, azt kapjuk, hogy

(18) $\quad (S_b g)(0) \cdot (S_b x)(0) = (S_b f)(0) + b^{N+1} t(N)$

Innen viszont következik, hogy

(19) $\quad b^{N+1} t(N) = c = \text{konstans} \quad N > \max(\mathrm{Supp}f, \mathrm{Supp}\ (g * x))$.

Legyen $N_0 > \max(\mathrm{Supp}f, \mathrm{Supp}\ (g * x))$ olyan nagy, hogy már

$$\frac{|c|}{b^{N_0+1}} < 1.$$

Akkor (19) következtében $|t(N_0)| = \dfrac{|c|}{b^{N_0+1}} < 1$. Mivel $|t(N_0)|$ egész, $|t(N_0)| < 1$ csak ugy teljesülhet, ha $t(N_0) = 0$. De akkor (19) miatt

$$0 = b^{N_0+1} t(N_0) = c$$

és igy (19)-re való tekintettel

(20) $\quad t(N) = 0, \quad$ ha $\quad N > \max(\mathrm{Supp} f, \mathrm{Supp}(g * x))$

De akkor (18)-ból a bizonyitandó (16) egyenlőség is következik, és mivel $(S_b x)(0)$ egész szám, az is következik (16)-ból, hogy $(S_b f)(0)$ osztható $(S_b g)(0)$-val.
Ezzel a tételt igazoltuk.

6.3. **Tétel.** *Legyen* $f,g \in I, (S_b f)(0) \geqslant 0, (S_b g)(0) > 0$. *Ha* $(S_b f)(0)$ *osztható* $(S_b g)(0)$*-val akkor létezik az* (1) *egyenletrendszernek a* (2) *feltételek melletti* $x(n), t(n)$ *megoldása. Ekkor* $x \in I_+$ *és*

$$(21) \qquad (S_b x)(0) = \frac{(S_b f)(0)}{(S_b g)(0)}$$

Bizonyitás. Legyen

$$(22) \qquad m = \frac{(S_b f)(0)}{(S_b g)(0)} \qquad \text{és}$$

legyen $\xi = \xi(n) \in I_+$ olyan, hogy

$$(23) \qquad m = \sum_{k=0}^{\mathrm{Supp}\xi} \xi(k) b^k = (S_b \xi)(0),$$

vagyis (23) az m számnak b alapu számrendszerben felirt alakja. Azt fogjuk igazolni, hogy $x = \xi$ a (2) feltételek mellett megoldása az (1) egyenletrendszernek.

Az, hogy $0 \leqslant \xi(n) < b \quad (n = 0,1,2, \ldots)$ rögtön következik (23)-ból.

Helyettesitsük be az (1) egyenletek mindegyikében x helyébe a ξ függvényt. Akkor $t(-1) = 0$ ismeretében a $t(0), t(1), \ldots$ értékek szukcesszive meghatározhatók a

$$(24) \qquad (g * \xi)(n) - bt(n) = f(n) - t(n-1) \qquad (n = 0,1,2, \ldots)$$

egyenletekből. A tétel bizonyitásához azt kell igazolni, hogy a $t(n)$ számok mind egész számok.

Legyen N olyan egész szám, amelyre

$$(25) \qquad N \geqslant \max(\mathrm{Supp} f, \mathrm{Supp}(g * \xi)) = N_0$$

teljesül. Adjuk össze a (24) alatti egyenlőségek mindkét oldalát 0-tól N-ig, akkor kapjuk, hogy

$$\sum_{n=0}^{N} \Big(\sum_{k=0}^{n} g(n-k)\xi(k) \Big) b^n - \sum_{n=0}^{N} t(n) b^{n+1} = \sum_{n=0}^{N} f(n) b^n - \sum_{n=0}^{N} t(n-1) b^n,$$

ahonnan (25) következtében kapjuk, hogy

$$(26) \qquad (S_b g)(0) \cdot (S_b \xi)(0) - t(N)b^{N+1} = (S_b f)(0),$$

ugyanis

$$\sum_{n=0}^{N} t(n-1)b^n = \sum_{n=1}^{N} t(n-1)b^n = \sum_{n=0}^{N-1} t(n)b^{n+1},$$

De akkor (22) és (23) figyelembe vételével (26)-ból kapjuk, hogy

$$(27) \qquad t(N) = 0.$$

Minden $N \geqslant N_0$ mellett. Azt kell még belátni, hogy ha $n \leqslant N_0$, akkor is egész szám lesz $t(n)$. Az állitást N_0-ra már beláttuk. Tegyük fel az indukciós bizonyitáshoz, hogy $t(n)$ egész szám, ha $0 < n \leqslant N_0$. Akkor (24)-ből következik, hogy

$$t(n-1) = f(n) + bt(n) - (g * \xi)(n)$$

is egész szám. Tehát minden $n = 0,1, \ldots$ mellett $t(n)$ egész szám, és fennáll (24) is és ezzel igazoltuk a tételt.

7. A faktorális táblázatot készitő program

A következőkben megmutatjuk, hogy a fenti elméleti meggondolások hogyan alkalmazhatók a gyakorlatban. Mivel $k!$ értéke rohamosan nő az k értékével nagy egész számokkal való számolásra adott módszerünk bemutatására faktoriális tábálzatot készitő ALGOL 60 programot irtunk.

Az eredményt egy fix felső korlátu w vektor tároloja. A vektor elemei az eredményt bináris kodolásu 10^d alapu ábrázolásban tartalmazzák. Az egyes faktoriálisok kiszámitásánál a vektor ténylegesen felhasznált hossza csak az értékes számjegyeket tartalmazó Supp w, amit az egymás utáni értékek kiszámitásánál a program határoz meg. Az eredményben gyorsan halmozódó számvégi nullákat a program csak megjegyzi, de nem számol velük. Ezt a Norm w eljárás oldja meg.

A $k!$-ból a $(k+1)!$ kiszámitása három lépésben történik:

I. A $k!$ jegyeit tartalmazó w vektort beszorozza $(k+1)$-gyel a hosszu egész számnak rövid egész számmal való szorzásra vonatkozó 2.3 tétel alapján.

II. Az igy kialakult w vektor ekvivalens $(k+1)!$-sal, de elemei $(k+1)!$-nak nem a 10^d alapu számrendszerbeli jegyeit tartalmazzák. A w vektort alávetjük a Tb transzformációnak, hogy a keletkezett vektorban megkapjuk $(k+1)!$-nak 10^d alapu számrendszerbeli jegyeit.

III. A Supp w és a Norm w eljárások alkalmazása.

A kiszámitott érték kiiratása egy előre megadott "kezd" értéktől, vagy egy kényszerités hatására történik.

A számolás során a 10^d alapszámot változtatjuk abból a célból, hogy az adott gép hardware sajátosságaiból adódó tulcsordulást kivédjük.

Esetünkben $k = 838$-ig 10^4, utána 10^3 az alapszám. Az adott programba a további csökkenést nem épitettük be, ezért 8388!-nál nagyobb, számot a program már nem számol helyesen (még akkor sem ha a w vektort a kellő hosszuságura kb. 10 000-re deklaráljuk).

A "RAJT" cimkéjü feltételes utasitás kis módositásával azonban a jelen program 847 344! értékét még ki tudja számitani a háttér tár felhasználásával. Ha ezen is túl akarunk menni, akkor másik programot kell irni, amelyben már két hosszú szám szorzatára kell eljárást készitení a konvolucios szorzás felhasználásával.

A programot ténylegesen 1003! kiszámitásáig futtattuk le. (1003! egy 2577 jegyü szám. Annak érzékeltetésére, hogy ez milyen nagy szám, legyen szabad megemliteni, hogy egy olyan sugarú gömbben, amely a földtől az Ursa Ma II. galaxishalmazig terjed, ami kb. félmilliárd fényév, legfeljebb annyi atom van, amelynek számjegyeinek száma 100-nál kevesebb.) A futási idő 10 perc 29 sec volt, amiből látszik, hogy egy szám faktoriálisának a kiszámitásra kb. 1/2 sec időre volt a gépnek szüksége átlagban (persze az elején kevesebb a végén több idő kellett).

Megjegyezzük, hogy ha assembler nyelv szintjéig mennénk le, vagy méginkább, ha gépi kódban 2^m alapu számrendszerben dolgoznánk, lényegesen gyorsabb lenne a program lefutása.

Ha a számitógép gyárak a fentiekben vázolt hosszú aritmetikát a gép alap-software-jébe épitenék be, akkor a jelenleg használatos programnyelvek némi módositásával még kis számitógépek esetén is tetszőlegesen előre adott pontossággal lehetne számolni viszonylag gyorsan.

Irodalom

[1] L. Berg,: Einführung in die Operatorenrechnung, Berlin 1962.

```
begin comment FAKT;
 integer korlaat;
 korlaat=9500;
 begin
  integer i,j,k,l,m,n,r,s,a,b,c,d,e,f,g,kezd,norm;
  Boolean nagy;
  integer array w[0:korlaat];
  procedure kiir(k,w,s,norm,c,d,f);
   integer k,s,norm,c,d,f;
   integer array w;
   begin
    integer i,r,l,g;
    r=10;
    for i=1 step 1 until d do
     if w[0]÷r×r≠w[0] then go to ki else r=r×10;
ki: if r>10 then
     begin r=r÷10;
      for i=0 step 1 until s do
       w[i]=w[i]÷r+(w[i+1]-w[i+1]÷r×r)×(c÷r);
      norm=norm+ln(r)/ln(10.0);
      if w[s]=0 then s=s-1;
     end pontos normirozaas;
    comment itt helyezkedik el a kiiroo programreesz;
   end kiir;
L0:k=s=norm=w[0]=0;
  copy(korlaat,w[0],w[1]);
  nagy=false; c=10000; d=4; f=60; e=f÷d; w[0]=1;
  comment egyeeb programkezdeti szervezeesek;
  kiir(k,w,s,norm,c,d,f);
RAJT:k=k+1;
  if k=838 then
rend:begin
       nagy=true; c=1000; d=3; e=f÷d;
       l=r=(s+1)/3+.2; s=3×r-1;
```

```
      for i:=s step -3 until 2 do
       begin m:=w[i+r]:=w[i]÷10; j:=w[i-1]÷100;
        w[i+r-1]:=(w[i]-m×10)×100+j; m:=w[i-2]÷1000;
        w[i+r-2]:=(w[i-1]-j×100)×10+m;
        w[i+r-3]:=w[i-2]-m×1000;
        r:=r-1;
       end i;
      s:=s+1;
     end bin. kod. 10^4-es aabr.-rool 10^3-asra aatteerees;
  b:=0;
  for i:=0 step 1 until s do
   begin l:=w[i]×k+b; b:=l÷c; w[i]:=l-b×c;
   end i;
ST:if b≠0 then
    begin l:=b÷c; w[i]:=b-l×c; b:=l; i:=i+1;
     go to ST;
    end 2.3 tetel es Tb transzformacio;
  s:=i-1; comment Supp w;
  i:=-1;
  for i:=i+1 while w[i]=0 do ;
  if i>0 then
   begin copy(s+1,w[i],w[0]); s:=s-i; norm:=norm+d×i;
   end durva normirozaas;
  kiir(k,w,s,norm,c,d,f);
  comment a befejezeest eloeiro felteetelek vizsgaalata;
  go to RAJT;
 end;
end FAKT;
```

Summary

Accurate calculation with arbitrary large integers by means of digital computers

E. Gesztelyi – P. Jékel

The paper deals with the theoretical background of an integer arithmetic applicable for accurate calculation with large integers independent of the specific hardware feature of the given digital computer.

A function is called arithmetical when it is defined on the set of nonnegative integers. The arithmetical function f is said to be an integral valued function if $f(n)$ is an integer for every n and such that $f(n) = 0$ for $n > N$ where the number N depends on f. In the set I of integral valued arithmetical functions let the addition be defined in the usual way and the multiplication as the convolution

$$(1) \qquad \sum_{k=0}^{n} f(n-k)g(k)$$

This way I becomes an integral domain.

Let $b \geqslant 0$ be a fixed integer that we consider the base of a number system. We define in I an equivalence relation (depending on b) as follows. We say that the functions $f,g \in I$ are equivalent with respect to b (written $f \overset{b}{\sim} g$) iff

$$(S_b f)(0) = (S_b g)(0)$$

where S_b is such a transformation that

$$(S_b f)(n) = \begin{cases} \sum_{k=0}^{N} f(k)b^k & \text{if} \quad n = 0 \\ 0 & \text{if} \quad n > 0 \end{cases}$$

This equivalence relation is compatible with respect to the addition and multiplication (defined by (1)). Thus we can construct the corresponding factor ring I/b in the usual way. The ring I/b is isomorphic to the ring of integers.

We define a transformation T_b as follows

$$s(0) = f(0)$$

$$s(n+1) = f(n+1) + [\frac{s(n)}{b}]$$

$$(T_b f)(n) = s(n) - b[\frac{s(n)}{b}]$$

where [ ] = entier ( ).

Let

$$I_+ = \{ f \mid f \epsilon I,\ 0 \leqslant f(n) < b,\ n = 0,1, \ldots \}$$

and

$$I_- = \{ g \mid b - 1 - g \epsilon I_+ \} .$$

Then the following statements are true.

1. If $f \epsilon I_+$ then $(T_b f)(n)$ is the $(n+1)$-th digit of $(S_b f)(0)$ with respect to the base b.
2. $f \overset{b}{\sim} g$ iff $T_b f = T_b g$ $(f,g \epsilon I)$ (Theorem 4.2.)
3. If $f \epsilon I$ and $(S_b f)(0) \geqslant 0$ then $T_b f \epsilon I_+$ and if $(S_b f)(0) < 0$ then $T_b f \epsilon I_-$ (Theorem 4.1.).
4. $T_b^2 = T_b$ (Theorem 4.3.).
5. Theorem 4.4.: $T_b(f + g) = T_b(T_b f + T_b g)$,
   $T_b(f * g) = T_b(T_b f * T_b g)$.
6. If $f \epsilon I_+ \cup I_-$ then $T_b f = f$.

If follows from the above facts that we are able to write programs for the computation of algebraic expressions consisting of integers even if they are very large. Let f and g be two integer procedures which generate the digits of $(S_b f)(0)$ and $(S_b g)(0)$, respectively, in the number system of base b. Since $f,g \epsilon I$, if we take the sum or the convolution of f and g we obtain a result h which is also in I. If we submit h to the procedure T_b then we get functions which are in I_+ or I_- depending on the sign of $(S_b h)(0)$.

Thus the values of $T_b h$ provide the digits of $(S_b h)(0)$ if $T_b h \epsilon I_+$ and if $T_b h \epsilon I_-$ then the complement of $T_b h$ gives the digits of the number $-(S_b h)(0)$.

We have also presented algorithms for the division.

To show the applicability of the theory we have written a program in ALGOL 60 for the computation of the factorials. We have flowed the program by the computer ODRA 1204 to calculate the factorials from 1 to 1003.

Р е з ю м е

Точные расчеты с произвольно большими числами на ЭВМ

Е. Гестельи – П. Йекель

В настоящей работе обсуждаются теоретические основы целой машинной арифметики, с помощью которой на любой ЭВМ, независимо от ее технического обеспечения, могут быть реализованы точные расчеты с произвольно большими целыми числами.

Функции определенные на множестве неотрицательных чисел назовем арифметическими функциями.

Арифметическую функцию назовем целой функцией если ее значения целые числа, за исключением, быть может, конечного числа значений функции равных нулю.

На множестве I целых функций определим обычным образом операцию сложения, а также операцию умножения со сверткой:

$$(f * g)(n) = \sum_{k=0}^{n} f(n-k)g(k) \qquad (f,g \in I) \tag{1}$$

Таким образом I область целостности.

Пусть $b \geq 2$ фиксированное целое число, которое принимаем за основание системы счисления.

Определим на I отношение эквивалентности:
будем говорить, что f и $g \in I$ эквивалентны относительно $f \overset{b}{\sim} g/$, если

$$\sum_{n=0}^{\infty} f(n)b^n = \sum_{n=0}^{\infty} g(n)b^n \tag{2}$$

отношение эквивалентности (2) совместимо на I.

Покажем, что соответствующая фактор-структура I/b изоморфна кольцу целых чисел.

Определим T_b – преобразование следующим образом:

Пусть f некоторая арифметическая функция, и пусть

$$\Delta(0) = f(0)$$
$$\Delta(n+1) = f(n+1) + \left[\frac{\Delta(n)}{b}\right]$$
$$(T_b f)(n) = s(n) - b\left[\frac{\Delta(n)}{b}\right],$$

где [ ] = entier ( ).

Обозначим через

$$I_+ = \{ f / f \epsilon I,\ 0 \leqslant f(n) < b, n = 0,1,2, \ldots \}$$

и через

$$I_- = \{ g / b - 1 - g \epsilon I_+ \} .$$

Тогда справедливы следующие утверждения

1/ Если $f \epsilon I_+$, тогда в системе счисления с основанием b $f(n)$ дает $n+1$ -ий знак числа $(S_b f)(0) = \sum_{n=0}^{N} f(n) b^n$

2/ $f \overset{b}{\sim} g$ в том и только в том случае, если $T_b f = T_b g$ /теорема 4.2/

3/ Если $f \epsilon I$ и $(S_b f)(0) \geqslant 0$, тогда $T_b f \epsilon I_+$, а также если $(S_b f)(0) < 0$, тогда $T_b f \epsilon I_-$ /теорема 4.1/

4/ $T_b^2 = T_b$ /теорема 4.3/

5/ $T_b(f + g) = T_b(T_b f + T_b g)$,

$T_b(f * g) = T_b(T_b f * T_b g), f, g \epsilon I$ /теорема 4.4/

6/ Если $f \epsilon I_+ \cup I_-$, тогда $T_b f = f$.

Из вышесказанного следует, что сумма и умножение со сверткой функций f и g, при помощи которых мы генерируем цифры целых чисел $(S_b f)(0)$ и $(S_b g)(0)$ в b -ричной системе счисления, принадлежат I .

Применяя к сумме или произведению (2) указаных функций преобразование T_b получаем функции лежащие в I_+ или в I_- , в зависимости от знака преобразования S_b .

Таким образом, мы получаем b-ричные цифры результата S_b непосредственно, если результат преобразования T_b лежит в I_+ , дополнение до $b-1$ дает нам цифры результата S_b со знаком минус, в том случае, если результат преобразования T_b лежит в I_- .

Программа на АЛГОЛе-60, написанная с использованием вышеописанного метода за эффективное время рассчитала, например, значение 1003! на ЭВМ ODRA 1204.