DISCRETE MATHEMATICS



Discrete Mathematics 142 (1995) 87-96

An abstract Möbius inversion formula with number-theoretic applications

Pentti Haukkanen

Department of Mathematical Sciences, University of Tampere, P.O. Box 607, SF-33101 Tampere, Finland

Received 5 August 1993

Abstract

An inversion formula for incidence functions is given. This formula is applied to certain types of number-theoretic identities, for example, to the arithmetical evaluation of Ramanujan's sum and to the identical equation of a class of multiplicative functions.

1. Introduction

It is well known that Ramanujan's sum C(m, n) has the arithmetical evaluation

$$C(m,n) = \sum_{d \mid (m,n)} d\mu(n/d).$$

Most generalized Ramanujan sums also have similar evaluations. For example, the unitary analogue of Ramanujan's sum $C^*(m, n)$ can be written as

$$C^*(m,n) = \sum_{\substack{d \mid m \\ d \parallel n}} d\mu^*(n/d)$$

On the other hand, it is well known that the divisor function σ_k satisfies the identity

$$\sigma_k(m)\sigma_k(n) = \sum_{d\mid (m,n)} \sigma_k(mn/d^2)d^k$$

An effort to understand better the above identity led to the study of the identity

$$f(m)f(n) = \sum_{d \mid (m,n)} f(mn/d^2)g(d)$$

The functions f satisfying this identity are said to be specially multiplicative. Some inverse forms of the above identities have also been studied in the literature. For further information on all these identities, see e.g. [5, Chs. 1 and 2]. It should be emphasized that Ramanujan discovered the arithmetical evaluation of C(m, n) [8] and an identity that may be referred to as an inverse form of the above identity for σ_0 [7].

0012-365X/95/\$09.50 © 1995—Elsevier Science B.V. All rights reserved SSDI 0012-365X(93) E 0209-M

The aim of this paper is

(i) to point out that the above identities, and also some other number-theoretic identities, have a common structure,

- (ii) to unify the treatment of inverse forms,
- (iii) to provide some new insight into this structure, and
- (iv) to develop an efficient tool for manipulations of this structure.

We present the common structure and its inverse form in an inversion theorem. This inversion theorem is given in Section 2. A lattice-theoretic approach provides a deeper understanding of this structure. The inversion theorem involves incidence functions of $(S \times S, \subseteq \leq)$ and (S, \leq) , where (S, \subseteq) and (S, \leq) are locally finite partially ordered sets. In the usual terminology this theorem may be called an abstract Möbius inversion formula (see e.g. [1, Ch. IV, 9]).

Section 3 introduces the concept of Narkiewicz's regular arithmetical convolution, which is needed in Sections 4–6.

In Section 4 the highly abstract inversion theorem of Section 2 is specialized to a more concrete form. The applications to Ramanujan's sum and to its generalizations and to specially multiplicative functions are also given there.

Section 5 provides a further example of the present structure in number theory. This example is an expression that counts the number of solutions of a restricted congruence.

In Section 6 the concepts of convolutions and principal functions are used in developing an efficient tool to manipulate the concrete form of the structure. This method makes it possible to present the structure in a very concise form and to interpret the inverse structure algebraically. It is pointed out that there are, unfortunately, some difficulties in generalizing this efficient method to the abstract setting of the structure.

2. An inversion theorem for incidence functions

Let (S, \leq) be a locally finite partially ordered set. Then a complex-valued function f on $S \times S$ is said to be an incidence function of (S, \leq) if f(x, y) = 0 unless $x \leq y$. The set of all incidence functions of (S, \leq) is denoted by $I(S, \leq)$. The convolution of f, g $(\in I(S, \leq))$ is defined by

$$(f \star g)(x, y) = \sum_{x \leqslant z \leqslant y} f(x, z)g(z, y).$$

The inverse of $f \in I(S, \leq)$ is defined by

$$f \star f^{-1} = f^{-1} \star f = \delta,$$

where $\delta \in I(S, \leq)$ is the identity function given by

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

It can be verified that the inverse of $f \in I(S, \leq)$ exists if and only if $f(x, x) \neq 0$ for all $x \in S$. The zeta function of $I(S, \leq)$ is defined by

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

The inverse of ζ is denoted by μ and is called the Möbius function of $I(S, \leq)$.

Let \subseteq be another locally finite partial order on S. Then $(S \times S) \subseteq \leq$ is a locally finite partially ordered set, where

 $(u, x) \subseteq \leq (v, y) \Leftrightarrow u \subseteq v, x \leq y.$

A complex-valued function f on $(S \times S) \times (S \times S)$ is said to be an incidence function of $(S \times S, \subseteq \leq)$ if f((u, x), (v, y)) = 0 unless $(u, x) \subseteq \leq (v, y)$. The set of all incidence functions of $(S \times S, \subseteq \leq)$ is denoted by $I(S \times S, \subseteq \leq)$.

We are now in a position to present the abstract Möbius inversion formula of this paper.

Theorem 1. Let (S, \subseteq) and (S, \leq) be locally finite partially ordered sets such that

$$x \leqslant y \implies x \subseteq y.$$

Suppose that $f, g \in I(S \times S) \subseteq g \in I(S \times S)$ and $h \in I(S, g \in S)$ with $h(x, x) \neq 0$ for all $x \in S$. Then

$$f((x, x), (v, y)) = \sum_{\substack{x \le z \le y \\ z \le v}} h(x, z)g((z, z), (v, y))$$
(1)

for all $v, x, y \in S$, if and only if

$$g((x, x), (v, y)) = \sum_{\substack{x \le z \le y \\ z \le v}} h^{-1}(x, z) f((z, z), (v, y))$$
(2)

for all $v, x, y \in S$, where h^{-1} is the inverse of h in $I(S, \leq)$.

Proof. Suppose that (1) holds. Then

$$\sum_{\substack{x \le z \le y \\ z \subseteq v}} h^{-1}(x, z) f((z, z), (v, y)) = \sum_{\substack{x \le z \le y \\ z \subseteq v}} h^{-1}(x, z) \sum_{\substack{z \le w \le y \\ w \subseteq v}} h(z, w) g((w, w), (v, y))$$
$$= \sum_{\substack{x \le w \le y \\ w \subseteq v}} \left(\sum_{\substack{x \le w \le y \\ w \subseteq v}} h^{-1}(x, z) h(z, w) \right) g((w, w), (v, y))$$
$$= \sum_{\substack{x \le w \le y \\ w \subseteq v}} \delta(x, w) g((w, w), (v, y))$$
$$= g((x, x), (v, y)).$$

Thus (2) holds. The converse is proved similarly. \Box

Corollary 1. Let (S, \subseteq) and (S, \leq) be locally finite partially ordered sets such that

 $x \leq y \Rightarrow x \subseteq y.$

Suppose that $f, g \in I(S \times S) \subseteq I(S \times S)$. Then

$$f((x, x), (v, y)) = \sum_{\substack{x \le z \le y \\ z \le v}} g((z, z), (v, y))$$

for all $v, x, y \in S$, if and only if

$$g((x, x), (v, y)) = \sum_{\substack{x \le z \le y \\ z \le v}} \mu(x, z) f((z, z), (v, y))$$

for all $v, x, y \in S$, where μ is the Möbius function of $I(S, \leq)$.

Corollary 2. Let (S, \leq) be a locally finite partially ordered set. Suppose that $f, g \in I(S, \leq)$. Then

$$f(x, y) = \sum_{x \leqslant z \leqslant y} g(z, y)$$

for all $x, y \in S$, if and only if

$$g(x, y) = \sum_{x \leq z \leq y} \mu(x, z) f(z, y)$$

for all $x, y \in S$, where μ is the Möbius function of $I(S, \leq)$.

Corollary 1 follows from Theorem 1 by taking $h = \zeta$, and Corollary 2 from Corollary 1 by taking v = y. Note that Corollary 2 is a form of the classical Möbius inversion formula for incidence functions.

3. Regular arithmetical convolutions

For each n, let A(n) be a subset of the set of positive divisors of n. The elements of A(n) are said to be the A-divisors of n. The A-convolution of two arithmetical functions f and g is defined by

$$(f \star_A g)(n) = \sum_{d \in A(n)} f(d)g(n/d).$$

Narkiewicz [6] defined an A-convolution to be regular if

(a) the set of arithmetical functions forms a commutative ring with unity with respect to the ordinary addition and the A-convolution;

(b) the A-convolution of multiplicative functions is multiplicative;

(c) the function $\equiv 1$ has an inverse μ_A with respect to the A-convolution, and $\mu_A(n) = 0$ or -1 whenever n is a prime power.

The inverse of an arithmetical function f such that $f(1) \neq 0$ with respect to the A-convolution is defined by

$$f \star_A f^{-1} = f^{-1} \star_A f = \delta,$$

where $\delta(1) = 1$ and $\delta(n) = 0$ for n > 1.

- It can be proved (see [6]) that an A-convolution is regular if and only if
- (i) $A(mn) = \{ de: d \in A(m), e \in A(n) \}$ whenever (m, n) = 1,
- (ii) for each prime power $p^a > 1$ there exists a divisor $t = \tau_A(p^a)$ of a such that

$$A(p^{a}) = \{1, p^{t}, p^{2t}, \dots, p^{rt}\},\$$

where rt = a, and

$$4(p^{it}) = \{1, p^{t}, p^{2t}, \dots, p^{it}\}, \quad 0 \le i < r.$$

For example, the Dirichlet convolution D, where D(n) is the set of all positive divisors of n, and the unitary convolution U, where

$$U(n) = \{d > 0: d \mid n, (d, n/d) = 1\} = \{d > 0: d \mid n\},\$$

are regular.

The following more specialized properties are needed in the applications. Here A always denotes a regular arithmetical convolution.

A positive integer n is said to be A-primitive if $A(n) = \{1, n\}$. The generalized Möbius function μ_A is the multiplicative function given by

$$\mu_A(p^a) = \begin{cases} -1 & \text{if } p^a (>1) \text{ is } A\text{-primitive,} \\ 0 & \text{if } p^a \text{ is non-}A\text{-primitive.} \end{cases}$$

In particular, one denotes $\mu_D = \mu$ and $\mu_U = \mu^*$.

The symbol $(m, n)_A$ denotes the greatest divisor of m which belongs to A(n). In particular, one denotes $(m, n)_D = (m, n)$ and $(m, n)_U = (m, n)_*$.

A positive integer *n* is said to be an *A*-square if for each $p^a || n$, *a* is of the form a = 2st, where $t = \tau_A(p^a)$. The generalized Liouville function λ_A is defined to be the multiplicative function given by $\lambda_A(p^a) = (-1)^b$, where a = bt, $t = \tau_A(p^a)$. Note that $\lambda_U = \mu^*$ and $\lambda_D = \lambda$, the classical Liouville function. (The connection between the concepts of *A*-square and λ_A is pointed out in [3, Example 2].)

Let k be a positive integer. A positive integer n is then said to be kth power free with respect to A-convolution if for each $p^a || n$, we have s < k, where a = st, $t = \tau_A(p^a)$. The generalized Möbius function $\mu_{A,k}$ is defined to be the multiplicative function given by $\mu_{A,k}(p^a) = -1$ if a = kt, $t = \tau_A(p^a)$, and = 0 otherwise. Note that $\mu_{D,k}$ is the Klee function, $\mu_{A,1} = \mu_A$, and $\mu_{U,k}$ (k > 1) is the identity function δ , where $\delta(1) = 1$ and $\delta(n) = 0$ for n > 1.

4. An inversion theorem for arithmetical functions of two variables

Let A and B be regular convolutions, and define the relation $\subseteq \leq$ on $\mathbb{N} \times \mathbb{N}$ by

 $(u, x) \subseteq \leq (v, y) \Leftrightarrow u \in B(v), x \in A(y).$

Let f(m, n) be an arithmetical function of two variables. Then we can associate with f an incidence function f' of $(\mathbb{N} \times \mathbb{N}, \subseteq \leq)$ defined by

$$f'((u, x), (v, y)) = f(v/u, y/x)$$

if $u \in B(v)$, $x \in A(y)$, and = 0 otherwise. It is easy to see that the mapping $f \to f'$ is one-one.

Theorem 1 can be specialized for arithmetical functions as follows.

Theorem 2. Let A and B be regular convolutions such that for each $n \in \mathbb{N}$, A(n) is a subset of B(n). Let h(n) be an arithmetical function with $h(1) \neq 0$. Then

$$f(m,n) = \sum_{\substack{d \in B(m) \\ d \in A(n)}} h(d)g(m/d, n/d)$$
(3)

for all $m, n \in \mathbb{N}$, if and only if

$$g(m,n) = \sum_{\substack{d \in B(m) \\ d \in \mathcal{A}(n)}} h^{-1}(d) f(m/d, n/d)$$
(4)

for all $m, n \in \mathbb{N}$, where h^{-1} is the inverse of h with respect to the A-convolution.

In the next examples we present applications of Theorem 2 to number-theoretic expressions. Throughout the examples B(m) = D(m), the set of all positive divisors of m.

Example 1. The generalized Ramanujan's sum $C_A(m, n)$ is defined by

$$C_A(m,n) = \sum_{\substack{x \pmod{n} \\ (x,n)_A = 1}} \exp(2\pi i m x/n)$$

(see [5, p. 164]). It is known that

$$C_A(m,n) = \sum_{\substack{d \in A(n) \\ d|m}} d\mu_A(n/d).$$
⁽⁵⁾

Thus $C_A(m, n)$ satisfies (3) with h(n) = n and $g(m, n) = \mu_A(n)$. Therefore application of (4) gives

$$\mu_A(n) = \sum_{\substack{d \in A(n) \\ d \mid m}} \mu_A(d) \, dC_A(m/d, n/d). \tag{6}$$

In particular, in the unitary case, we obtain

$$\mu^{*}(n) = \sum_{\substack{d \mid n \\ d \mid m}} \mu^{*}(d) \, dC^{*}(m/d, n/d), \tag{7}$$

where $C^*(m, n)$ is the usual notation for the unitary analogue of Ramanujan's sum. In the case of the Dirichlet convolution (6) reduces to

$$\mu(n) = \sum_{d \mid (m, n)} \mu(d) \, dC(m/d, n/d), \tag{8}$$

where C(m, n) is the classical Ramanujan sum. It should be noted that Ramanujan discovered the structure of (5) in the fundamental case A = D in his celebrated paper [8].

Example 2. Let

$$\beta_A(m,n) = \sum_x \exp(2\pi i m x/n)$$

where the sum is over all integers x such that $x \pmod{n}$ and $(x, n)_A$ is an A-square. It can be proved that

$$\beta_A(m,n) = \sum_{\substack{d \in A(n) \\ d|m}} d\lambda_A(n/d)$$
(9)

(cf. [3, Theorem 9]). Thus $\beta_A(m, n)$ satisfies (3) with h(n) = n and $g(m, n) = \lambda_A(n)$. Therefore application of (4) gives

$$\lambda_A(n) = \sum_{\substack{d \in A(n) \\ d \mid m}} \mu_A(d) \, d\beta_A(m/d, n/d). \tag{10}$$

In the unitary case, (10) reduces to (7), and in the case of the Dirichlet convolution, (10) reduces to

$$\lambda(n) = \sum_{d \mid (m,n)} \mu(d) d\beta(m/d, n/d).$$
(11)

Note that β is the function B of Sivaramakrishnan [10, p. 202].

Example 3. Let

$$C_{A,k}(m,n) = \sum_{x} \exp(2\pi i m x/n),$$

where the sum is over all integers x such that $x \pmod{n}$ and $(x, n)_A$ is kth power free with respect to A-convolution. It can be proved that

$$C_{A,k}(m,n) = \sum_{\substack{d \in A(n) \\ d|m}} d\mu_{A,k}(n/d)$$
(12)

(cf. [3, Theorem 9]). Thus $C_{A,k}(m,n)$ satisfies (3) with h(n) = n and $g(m,n) = \mu_{A,k}(n)$. Therefore application of (4) gives

$$\mu_{A,k}(n) = \sum_{\substack{d \in A(n) \\ d|m}} \mu(d) \, dC_{A,k}(m/d, n/d) \,. \tag{13}$$

Example 4. An arithmetical function f is said to be specially multiplicative [4] if there exists a completely multiplicative function g such that

$$f(m)f(n) = \sum_{d|(m,n)} f(mn/d^2)g(d)$$
(14)

for all $m, n \in \mathbb{N}$. Thus f(m)f(n) satisfies (3) with h(n) = g(n) and g(m, n) = f(mn). Therefore application of (4) gives the following well-known identity

$$f(mn) = \sum_{d \mid (m,n)} f(m/d) f(n/d) \mu(d) g(d)$$
(15)

for all $m, n \in \mathbb{N}$ (see e.g. [5, p. 19]). Equations of the types (14) and (15) are also termed Busche-Ramanujan equations, since the study of these equations arose from the observations of Busche [2] and Ramanujan [7] that the divisor functions σ_k possess this structure (see also [5, p. 25]).

5. A congruence

Theorem 3. Let $a_1, a_2, ..., a_s$ be integers such that $((a_j), n) = 1$, where (a_j) is the g.c.d. of $a_1, a_2, ..., a_s$. Let Q(m, n, s) denote the number of solutions $\langle x_1, x_2, ..., x_s \rangle$ of the congruence

$$m \equiv a_1 x_1 + a_2 x_2 + \dots + a_s x_s \pmod{n} \tag{16}$$

such that $((x_j), n)_A = 1$. Then

$$Q(m, n, s) = \sum_{\substack{d \in \mathcal{A}(n) \\ d \mid m}} \mu_{\mathcal{A}}(d) (n/d)^{s-1}$$
(17)

and

$$n^{s-1} = \sum_{\substack{d \in A(n) \\ d|m}} Q(m/d, n/d, s).$$
(18)

Proof. Let $p_i^{t_i}$, i = 1, 2, ..., r, be the A-primitive prime powers such that $p_i^{t_i} \in A(n)$, $p_i^{t_i}|m$, and let N_i , i = 1, 2, ..., r, be the set of solutions of the congruence such that $p_i^{t_i}|(x_j)$, where (x_j) is the g.c.d. of $x_1, ..., x_s$. Then Q(m, n, s) is the number of solutions of (16) which do not belong to $N_1 \cup \cdots \cup N_r$.

If $1 \le i_1 < \cdots < i_k \le r$, then $card(N_{i_1} \cap \cdots \cap N_{i_k})$ is the number of unrestricted solutions of

$$m/(p_{i_1}^{t_{i_1}}\cdots p_{i_k}^{t_{i_k}}) \equiv a_1y_1 + \cdots + a_sy_s (\mod n/(p_{i_1}^{t_{i_1}}\cdots p_{i_k}^{t_{i_k}})).$$

Since $((a_i), n/(p_{i_1}^{t_{i_1}} \cdots p_{i_k}^{t_{i_k}})) = 1$, then

 $\operatorname{card}(N_{i_1} \cap \cdots \cap N_{i_k}) = [n/(p_{i_1}^{t_{i_1}} \cdots p_{i_k}^{t_{i_k}})]^{s-1}$

(see [5, Proposition 3.1]). Thus, by the inclusion-exclusion principle (see [5, p. 12]),

$$Q(m, n, s) = n^{s-1} + \sum_{k=1}^{r} (-1)^{k} \sum_{\substack{1 \le i_{1} \le \dots \le i_{k} \le r \\ d \mid m}} [n/(p_{i_{1}}^{t_{i_{1}}} \cdots p_{i_{k}}^{t_{i_{k}}})]^{s-1}$$

$$= \sum_{\substack{d \in A(n) \\ d \mid m}} \mu_{A}(d)(n/d)^{s-1}.$$

This proves (17). Further, application of Theorem 2 gives (18). \Box

Remark. Eq. (17) is known in the case A = D (see [5, p. 127]).

6. Principal functions

For an arithmetical function h(n), Vaidyanathaswamy [11] defined the principal function $P_h(m, n)$ equivalent to h as follows:

$$P_h(m,n) = \begin{cases} h(n) & \text{if } m = n, \\ 0 & \text{otherwise.} \end{cases}$$

Vaidyanathaswamy also connected this concept to the theory of the Dirichlet convolution. Here we show that principal functions connected to regular convolutions is a very efficient tool in the context of the number-theoretic structure of this paper.

Let A and B be regular convolutions such that for each $n \in \mathbb{N}$, A(n) is a subset of B(n). The BA-convolution of arithmetical functions f(m, n) and g(m, n) is then defined by

$$(f \star_{BA} g)(m,n) = \sum_{d \in B(m)} \sum_{e \in A(n)} f(d,e)g(m/d,n/e).$$

The identity Δ under the *BA*-convolution is given by $\Delta(m, n) = 1$ if m = n = 1, and = 0 otherwise. It can be verified that

$$P_{h_1} \star_{BA} P_{h_2} = P_{(h_1 \star_A h_2)}, \qquad P_{\delta} = \Delta, \qquad (P_h)^{-1} = P_{h^{-1}},$$

where $(P_h)^{-1}$ is the inverse of P_h with respect to the *BA*-convolution and h^{-1} is the inverse of *h* with respect to the *A*-convolution. Therefore, by simple algebraic manipulations, we obtain

$$f = (P_h) \star_{BA} g \iff g = (P_{h^{-1}}) \star_{BA} f,$$

which is Theorem 2 in a very concise form.

The question naturally arises whether this efficient tool can be extended for use with the incidence functions. Unfortunately, we encounter difficulties, which we point out here. For $h \in I(S, \leq)$ we define the principal function $P_h \in I(S \times S, \leq \leq)$ by

$$P_h((u, x), (v, y)) = \begin{cases} h(u, v) & \text{if } u = x, v = y, \\ 0 & \text{otherwise.} \end{cases}$$

The convolution of f and $g \in I(S \times S, \subseteq \leq)$ is defined by

$$(f \circ g)((u, x), (v, y)) = \sum_{\substack{u \leq w \leq v \\ x \leq z \leq y}} f((u, x), (w, z))g((w, z), (v, y)).$$

It can be verified that under the assumptions of Theorem 1,

$$P_{h_1} \circ P_{h_2} = P_{(h_1 \star h_2)},$$

where \star is the convolution in $I(S, \leq)$. Unfortunately, however,

$$P_{\delta} \neq \Delta$$
,

where Δ is the identity under the convolution in $I(S \times S) \subseteq \langle \rangle$. To be more precise, P_{δ} has the property

$$(P_{\delta} \circ f)((u, x), (v, y)) = \begin{cases} f((x, x), (v, y)) & \text{if } u = x, \\ 0 & \text{otherwise} \end{cases}$$

for all $f \in I(S \times S) \subseteq \langle \rangle$, while Δ has the property

 $\Delta \circ f = f \circ \Delta = f$

for all $f \in I(S \times S) \subseteq I(S \times S)$, $\subseteq I(S \times S)$, $I(S \times S)$,

References

- [1] M. Aigner, Combinatorial Theory (Springer, New York, 1979).
- [2] E. Busche, Lösung einer Aufgabe über Teileranzahlen, Mitt. Math. Gesell. Hamburg 4 (1906) 229-237.
- [3] P. Haukkanen, Some generalized totient functions, Math. Student 56 (1988) 65-74.
- [4] D.H. Lehmer, Some functions of Ramanujan, Math. Student 27 (1959) 105-116.
- [5] P.J. McCarthy, Introduction to Arithmetical Functions (Springer, New York, 1986).
- [6] W. Narkiewicz, On a class of arithmetical convolutions, Colloq. Math. 10 (1963) 81-94.
- [7] S. Ramanujan, Some formulae in the analytic theory of numbers, Mess. Math. 45 (1916) 81-84.
- [8] S. Ramanujan, On certain trigonometrical sums and their applications in the theory of numbers, Trans. Cambridge Philos. Soc. 22 (1918) 259-276.
- [9] C.-G. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, Z. Wahr. 2 (1964) 340-368.
- [10] R. Sivaramakrishnan, Classical theory of arithmetic functions, in: Monographs and Textbooks in Pure and Applied Mathematics, Vol. 126 (Marcel Dekker, New York, 1989).
- [11] R. Vaidyanathaswamy, The theory of multiplicative arithmetic functions, Trans. Amer. Math. Soc. 33 (1931) 579-662.