Commun. Math. Phys. 127, 239–252 (1990)

Communications in Mathematical Physics

© Springer-Verlag 1990

Supersymmetry and the Möbius Inversion Function

Donald Spector

Institute for Theoretical Physics, University of Utrecht, Princetonplein 5, Postbus 80.006, NL-3508 TA Utrecht, The Netherlands

Abstract. We show that the Möbius inversion function of number theory can be interpreted as the operator $(-1)^F$ in quantum field theory. Consequently, we are able to provide physical interpretations for various properties of the Möbius inversion function. These include a physical understanding of the Möbius Inversion Formula and of a result that is equivalent to the prime number theorem. Supersymmetry and the Witten index play a central rôle in these constructions.

1. Introduction

One of the most fundamental functions of number theory is the Möbius inversion function μ [3]. It is a function whose domain is the positive integers, and which is defined as follows. Let us say that an integer is *squarefree* if it is divisible by no perfect square other than 1, which of course means that in the prime decomposition of a squarefree number, each prime factor present appears exactly once. Then the function μ is defined by

$$\mu(n)=\begin{cases}+1 & \text{if } \; n \text{ is squarefree with an even number of prime factors}\\ -1 & \text{if } \; n \text{ is squarefree with an odd number of prime factors}\\ 0 & \text{otherwise.}\end{cases}$$

The function μ appears throughout number theory. For example, it plays a central rôle in the theory of Dirichlet convolution, and arises as well in several proofs of the prime number theorem. We will see in this paper that the function μ has a very natural physical interpretation. In the proper context, it is equivalent to $(-1)^F$, the operator that distinguishes fermionic from bosonic states and operators, with the fact that $\mu(n)=0$ when n is not squarefree being equivalent to the Pauli exclusion principle.

In this paper, we develop this identification between μ and $(-1)^F$. In so doing, we will be able to use physical arguments to derive and understand some of the fundamental results of arithmetic number theory. We will see in particular that supersymmetry is a key ingredient in understanding these results, and a special rôle will be played by the Witten index [24].

Developing a physical picture for number theoretic ideas is of importance for a variety of reasons. First, we find it very illuminating to uncover a rôle for the Witten index in number theory. We already know that the Witten index can help us understand various topological results. In non-linear sigma models, the Witten index computes the Euler characteristic of the target manifold, and other topological quantities (e.g., the Lefschetz number) are closely related [24]. More profoundly, the Witten index can serve as the principal tool in proving the Atiyah–Singer index theorem [2]. The Witten index itself can be viewed as a winding number of the Nicolai map [16] in function space [9]. That the Witten index should have a number theoretic interpretation, too, we feel is very interesting.

Elaborating the connection between topological and geometrical ideas with physical ones is a familiar theme in current research. Whether, for example, in attempts to obtain realistic string compactifications from Calabi–Yau manifolds [8] or the use of BRST invariance to obtain such quantities as the Donaldson invariants [25], the deep connections relating physics to geometry and topology are being actively explored. And, of course, group theory and algebraic results have very important and well-known physical manifestations, from crystallography to grand unification [12, 20].

We feel it is worthwhile to broaden this effort, and uncover some of the connections between number theory and physics, given how productive the interplay between physics and other areas of mathematics has been. Indeed, one can argue that any important mathematical structure ought to have a natural physical representation, a point of view which underlies much current work [26]; the work presented below helps to confirm that sentiment.

Furthermore, there are direct and indirect indications that number theory should play a rôle in string theory. At the most obvious level, modular functions arise in a key way in both string theory [13] and number theory [4, 18], though more abstract connections have been suggested. However, before one dives head in to the depths of number theory and strings, we think it is illuminating and instructive to map out some of the identifications one can draw between physics and number theory in an arena where everything is clearly- and well-understood. The number theory we address in this paper is thus perhaps not always as sophisticated as some of the modular function theory one can exploit in string theory. But since, by and large, our understanding of the connection between number theory and physics is limited, the results in this paper may serve as a foundation for further work, which may reduce the necessity of making inspired guesses or hopeful speculations in the future.

In addition, string theory seems to be an area of physics that lies at the junction of many areas of mathematics—number theory, complex analysis, group theory, geometry, and topology. A simpler "model" of such a junction now seems to be supersymmetry and the Witten index, which, as we see in this paper, have number theoretic as well as group theoretic, topological, and analytic (via the index theorem for differential operators) significance. It is worth learning whatever we can about how such a junction can occur.

Finally, it is clear that new ways of thinking about a subject often lead to new insights. The work in this paper enables us to use (supersymmetric) quantum field theory as a way of understanding number theory, and to use number theory as a way to understand quantum mechanics. Calculations and results that are natural in one language turn out to correspond to calculations and results that are natural in the other language. By establishing a unity between portions of these two literatures, we are hopeful that new insights will be forthcoming from these new ways of thinking.

The outline of the paper is as follows. First, we will suggest a natural way to associate the natural numbers with the states of a quantum system, and from this there immediately follows a correspondence between the operators in the quantum theory and arithmetic functions. With this set up, we will see that the Möbius inversion function μ arise from considering fermionic systems. We then proceed to derive some of the essential properties of μ. What is most interesting is that we use supersymmetry and the calculation of the Witten index in a class of supersymmetric theories to derive these properties. We then look at infinite sums involving μ, and obtain some important number theoretic results, again via the Witten index. One of the results we obtain is equivalent to the prime number theorem, one of the central achievements of number theory, in which the asymptotic density of prime numbers is computed. We also show that the particular quantum system we study when looking at these infinite sums is especially natural for expressing number theoretic quantities.

Following that, we derive the Möbius Inversion Formula, a result central to the study of arithmetic functions, and related to the possible factorizations of integers. This formula is not so hard in principle to derive mathematically, but in that context there is no especially clear picture of its meaning. However, this formula is very easy to understand in a physical context. More importantly, once we re-phrase the problem in a physical language, we will see that the answer (the Möbius Inversion Formula) is immediately suggested by the physics, and that this answer involves supersymmetry. Of course, Möbius himself knew nothing about supersymmetry, so it is quite compelling that such a different subject should very naturally yield some of his results.

We end the paper with some thoughts about what questions one might address next.

2. Gödel, Möbius, Witten

The key to our connection between number theory and quantum physics will be our (unconventional) choice for labeling the states of a theory. To motivate this choice, let us imagine we have a theory whose Hilbert space has as a basis the Fock space generated by a set of creation operators $b_1^\dagger, b_2^\dagger, b_3^\dagger, \ldots$. Thus the states of the theory are

$$\left(\prod_{j=1}^{r} (b_{i_j}^\dagger)^{\alpha_j}\right)|\text{vacuum}\rangle, \tag{2.1}$$

where i_j and α_j are positive integers. There is a countable infinity of creation operators and of states. Each state of the theory is determined uniquely by which creation operators act on the vacuum, and by how many times each one acts.

The positive integers possess a similar representation. The fundamental theorem of arithmetic tells us that each positive integer has a unique representation given by prime factorization. A number is specified by giving its prime factors, and by giving the maximum power of each prime factor that divides the number in question.

This suggests the following labelling of states. Let us represent the prime numbers in increasing sequence by $p_1, p_2, p_3, \ldots$. Thus, $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc. To each certain operator $b_i^\dagger$, we associate the prime number p_i. Then we label states as follows:

$$\left(\prod_{j=1}^{r} (b_{i_j}^\dagger)^{\alpha_j}\right)|\text{vacuum}\rangle \equiv |N\rangle, \tag{2.2}$$

where the number N is defined to be

$$N \equiv (p_{i_1})^{\alpha_1} \cdot (p_{i_2})^{\alpha_2} \cdot \cdots \cdot (p_{i_r})^{\alpha_r}. \tag{2.3}$$

Thus we have identified the unique "factorization" of a state into creation operators acting on a vacuum with the unique factorization of an integer into prime numbers. Consequently the state $|1\rangle$ is the vacuum; $|2\rangle$ and $|17\rangle$ are one-particle states; $|60\rangle$ is a four-particle state; etc. Clearly, the states $|N\rangle$, $N = 1, 2, 3, \ldots$, form a basis for the full Hilbert space of the theory. (We take the states $|N\rangle$ and analogous states throughout this paper to be normalized.)

This way of assigning a number to each state is unconventional in the physics literature, but it is not without precedent in general. In the mathematics literature, such a labeling is called a *Gödel numbering*, and is a central tool in establishing the absolutely fundamental Incompleteness Theorem [10]. We will repeatedly use this scheme for identifying states, as it will naturally lead us toward various number theoretic results.

For our purposes, it is only necessary to consider the following simple types of systems. First, take an ordinary quantum mechanical Hamiltonian H_0 with eigenstates Ψ_k of energy E_k. Let $b_k^\dagger$, be the creation operator asociated with the state Ψ_k upon second quantization. We now consider the field theory Hamiltonian H whose Hilbert space is the Fock space of states given by the above creation operators, and we only include the very limited additional interactions such that these Fock states $|N\rangle$ are in fact the eigenstates of the Hamiltonian H.

Having motivated our iosmorphism between states and integers in purely bosonic context above, we proceed to the substantive results of this paper. We look now at a supersymmetric field theory. As before, we have bosonic creation operators $b_1^\dagger, b_2^\dagger, b_3^\dagger, \ldots$, but now we have in addition fermionic creation operators $f_1^\dagger, f_2^\dagger, f_3^\dagger, \ldots$. The creation operators $b_k^\dagger$ and $f_k^\dagger$ are superpartners. Because we take the creation operators to create the eigenstates of the Hamiltonian, the supercharges interchange $b_k^\dagger$ and $f_k^\dagger$, without mixing in other operators. The states in our theory are

$$\left(\prod_{j=1}^{r} (b_{i_j}^\dagger)^{\alpha_j} (f_{i_j}^\dagger)^{\beta_j}\right)|\text{vacuum}\rangle \equiv |N, d\rangle, \tag{2.4}$$

where

$$N \equiv (p_{i_1})^{\alpha_1+\beta_1} \cdot (p_{i_2})^{\alpha_2+\beta_2} \cdot \ldots \cdot (p_{i_r})^{\alpha_r+\beta_r} \tag{2.5a}$$

and

$$d = (p_{i_1})^{\beta_1} \cdot (p_{i_2})^{\beta_2} \cdot \ldots \cdot (p_{i_r})^{\beta_r}. \tag{2.5b}$$

As before, we are using p_k to denote the k^{th} prime number. The number N indicates the "total excitation" of the state above the vacuum; the number d indicates which part of this excitation is fermionic. We note the following properties of the numbers α_j, β_j, d, and N. First, the numbers α_j can assume the values of any non-negative integer, while each of the β_j must be either 0 or 1, due to Fermi-Dirac statistics. Consequently, the number N can be any positive integer. On the other hand, d must be a squarefree integer. Furthermore, in any state $|N, d\rangle$, d must be a factor of N, a property we denote by "$d|N$". Finally, the above restrictions are the only ones, so that a list of all $|N, d\rangle$, where $d|N$ and d is squarefree is a list of a complete set of states of the theory.

For simplicity, not only do we restrict ourselves to the case that the states $|N, d\rangle$ are eigenstates of the field theory Hamiltonian, but we also consider only the case in which quantum mechanics Hamiltonian H_0 which gave us the creation operators has a purely discrete spectrum. One consequence of this is that the field theory has a mass gap and a discrete spectrum itself. For convenience, we can order the creation operators in a natural way, so that, for example, the state $|p, 1\rangle$ has a greater energy than the state $|\tilde{p}, 1\rangle$ when p and $\tilde{p}$ are prime numbers with $p > \tilde{p}$, but this is not necessary.

We now examine the Witten index of the above supersymmetric theory, that is, the number of bosonic minus the number of fermionic zero energy states. As is well-known, the Witten index is a topological invariant of a supersymmetric theory, unchanged by deformations in the parameters of the theory; it serves as a tool in analyzing the possibility of spontaneous supersymmetry breaking, dynamical or otherwise [24]. Since the theory has a discrete spectrum, the Witten index Δ can be computed reliably via the expression [9, 1]

$$\Delta = \operatorname{tr}[(-1)^F e^{-\beta H}], \tag{2.6}$$

and this expression will be independent of β, despite naïve appearances. Second, it is apparent in this case from an analysis of the Fock space that there is a unique vacuum state of zero energy, and consequently, on physical grounds, we see that the Witten index $\Delta = 1$. It is enlightening, however, to set up the computation of the Witten index using the notation developed above, (2.4) and (2.5).

Using (2.6), we have

$$\Delta = \sum_{N=1}^{\infty} \left(\sum_{d}{}' \langle N, d | (-1)^F e^{-\beta H} | N, d \rangle \right), \tag{2.7}$$

where the $\sum'$ means that we sum only over allowed values of d: those that are squarefree and divide N. The states $|N, d\rangle$ have two important properties for our purposes. First, the energy of (expectation value of H in) the state $|N, d\rangle$ is a function only of N and is independent of d. Second, the statistics (value of $(-1)^F$)

of the state $|N, d\rangle$ is a function only of d, and is independent of N, as it only depends on how many $f_k^\dagger$'s have acted.

We define $E_N = \langle N, d | H | N, d \rangle$. So

$$\Delta = \sum_{N=1}^{\infty} e^{-\beta E_N} \left(\sum_d{}' \langle N, d | (-1)^F | N, d \rangle \right). \tag{2.8}$$

Remember that despite appearances $\langle N, d | (-1)^F | N, d \rangle$ is independent of N. We can include non-squarefree d in the sum, provided we project them out. Thus for any positive integer d it is natural to define a function

$$\mu(d) = \begin{cases} \langle N, d | (-1)^F | N, d \rangle & \text{for} \quad d \text{ squarefree} \\ 0 \quad \text{otherwise} \end{cases} \tag{2.9}$$

For the squarefree values of d, $(-1)^F$ is equal to $+1$ or -1 when d has an even or odd number of prime factors, i.e., an even or odd number of fermionic excitations, respectively. Consequently this function μ defined in (2.9) is given by

$$\mu(d) = \begin{cases} +1 & \text{if} \quad d \text{ is squarefree with an even number of prime factors} \\ -1 & \text{if} \quad d \text{ is squarefree with an odd number of prime factors} \\ 0 & \text{otherwise.} \end{cases}$$

In other words, μ is exactly the Möbius inversion function advertised in the first paragraph of this paper. We see now that it corresponds exactly to the operator $(-1)^F$. Not only is the identification in itself interesting, it means that we now have physical tools available to analyze the properties of this function and its rôle in number theory.

We continue to study the Witten index to derive a key property of μ. We have

$$\Delta = \sum_{N=1}^{\infty} \left(e^{-\beta E_N} \sum_{d|N} \mu(d) \right), \tag{2.10}$$

where "$\sum_{d|N}$" means that we sum over all d that are factors of N. We know that $E_1 = 0$, and that $E_N > 0$ for $N > 1$, as this is a system with unbroken supersymmetry ($\Delta = 1$) and a mass gap. Furthermore, the set of states with a given value of $N > 1$ forms a complete bose-fermi symmetric representation of the supersymmetry algebra, as these states just constitute all ways of getting the same overall excitation with either bosonic or fermionic creation operators. Thus, these states give a net zero contribution to the Witten index, and we see that

$$\sum_{d|N} \mu(d) = 0 \quad \text{for} \quad N > 1. \tag{2.11}$$

On the other hand, since $\Delta = 1$ (or since the zero-energy ground state forms the lone singlet representation of the supersymmetry algebra in the Fock space), we have

$$\sum_{d|N} \mu(d) = 1 \quad \text{for} \quad N = 1. \tag{2.12}$$

These last two equations characterize perhaps the essential property of μ. Indeed, (2.11) and (2.12) are often taken as defining the function μ, with our earlier definition being derived from these properties. (It is these formulas one uses, also, in extending the notion of the Möbius inversion function to the study of groups and subgroups.)

We mention an alternative computation of the above results. Suppose all the E_N are non-degenerate, which certainly occurs in many systems. Then the fact that Δ is independent of β implies that $\sum_{d|N} [\mu(d)] = 0$ when $E_N > 0$, i.e., when $N > 1$, as otherwise the apparent β dependence of Δ in (2.6) would not be removed. Then the normalization that $\Delta = 1$ gives (2.12). This argument relies less explicitly on the details of the supersymmetry algebra and its representations in the Fock space, and is perhaps a somewhat more elegant argument than the first one given above. But the first argument, which spells out the connection between the factorizations of N with the different states in a representation of the supersymmetry algebra is quite instructive.

3. Infinite Sums and the Prime Number Theorem

We now wish to study infinite sums involving the function μ, and we do so by considering a particular theory. We continue to study the kind of supersymmetric field theory considered in the previous section, but we now restrict to the case that the state $|N,d\rangle$ has energy $E_N = \omega \log N$. The important feature of this choice is the energy is totally additive, and so we have a genuinely free theory, with all the states built of non-interacting excitations of energy $\omega \log p$. This is stronger than the statement that the energy only depends on N and not on d; we have further results such as that the energy of $|N,d\rangle$ is equal to the energy of the fermionic piece $|d,d\rangle$ plus the energy of the bosonic piece $|N/d,1\rangle$.

Consequently, since there is no interaction between the quanta associated with the various creation operators, we can write the Hamiltonian for this field theory as $H = H_b + H_f$, where $H_b(H_f)$ is the purely bosonic (fermionic) piece of the Hamiltonian. Since H_b and H_f are entirely non-interacting, we have that the Witten index Δ is given by

$$\Delta = \mathrm{tr}\,[(-1)^F e^{-\beta(H_b+H_f)}] = \mathrm{tr}\,[e^{-\beta H_b}]\cdot \mathrm{tr}\,[(-1)^F e^{-\beta H_f}]. \tag{3.1}$$

The fact that $\Delta = 1$ we know from the previous section. From the separation of the Hamiltonian, we see that

$$1 = \sum_{N=1}^{\infty} \langle N,1|e^{-\beta H_b}|N,1\rangle \cdot \sum_{d=1}^{\infty} \langle d,d|(-1)^F e^{-\beta H_f}|d,d\rangle. \tag{3.2}$$

Now $E_N = \omega \log N$, so, defining $s = \exp(\beta\omega)$, (3.2) becomes

$$1 = \sum_{N=1}^{\infty} \frac{1}{N^s} \cdot \sum_{d=1}^{\infty} \frac{\mu(d)}{d^s}. \tag{3.3}$$

The first term in the product comes from the bosonic Hamiltonian, and the second term from the fermionic Hamiltonian. The piece from the bosonic Hamiltonian is simply the Riemann zeta function, and converges for $s = \beta\omega > 1$. (The temperature $\beta^{-1} = \omega$ is essentially the Hagedorn temperature [14] of this theory.)

We therefore obtain the result that

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d^s} = \frac{1}{\zeta(s)}, \quad s > 1. \tag{3.4}$$

The identity (3.4), despite its physical origin, enables one to compute, for example, that the asymptotic density of squarefree integers is $(6/\pi^2)$ [17], a quantity which we can interpret now as relating to the ratio of two densities of states.

Suppose we consider the case $s = 1$ on the left-hand side of (3.4). For any $s = 1 + \varepsilon$ for arbitrarily small positive ε this formula is well-defined. In the limit that $\varepsilon \to 0$, the function $\zeta(1 + \varepsilon)$ diverges. Consequently, we conclude that

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d} = 0. \tag{3.5}$$

Strictly speaking, we have shown that if the sum on the left converges, then it converges to zero. It is a simple exercise to show it converges, going through the technicality of considering the partial sums of the series. The formula (3.5), which says that the partition function for the fermionic part of the theory vanishes at $\beta\omega = 1$, has a much deeper significance theory. It can be shown that this formula is equivalent to the prime number theorem [5], one of the central achievements of number theory. The prime number theorem determines the asymptotic density of the function $\pi(x)$, where $\pi(x)$ is equal to the number of numbers less than or equal to x that are prime; the result is that [3]

$$\pi(x) \sim \frac{x}{\log x}. \tag{3.6}$$

Thus we have the remarkable and surprising result that the Witten index calculation of a particular supersymmetric quantum system can be used to derive a result equivalent to the prime number theorem.

For $s = \beta\omega < 1$, it is harder to obtain any results. There is no obvious limiting procedure that will let us reach this domain for s. (One might invoke analytic continuation, but this has its own subtleties.) Indeed, the sum $\sum_{m=1}^{\infty} (\mu(m)/m^s)$ diverges for $s < 1$, but it is an open question in mathematics as to how fast such a sum diverges. The most interesting case is $s = 0$. It is Merten's Conjecture that the sum $\sum_{m=1}^{x} \mu(m)$ is asymptotic to $\sqrt{x}$ for large x. It would be especially interesting if we could use physical insight to address this question, as proving the conjectured asymptotic behavior of this sum would be equivalent to proving the Riemann hypothesis (and even a more limited result would shed light on the Riemann hypothesis) [21].

We would note one simple corollary of our above calculation. Let M be a totally multiplicative function over the natural numbers, i.e., suppose $M(z_1 z_2) = M(z_1)M(z_2)$ for all positive integers z_1 and z_2. Then if

$$\sum_{z=1}^{\infty} M(z) = \tilde{M}, \tag{3.7}$$

the sum

$$\sum_{z=1}^{\infty} \mu(z)M(z) = \tilde{M}^{-1}, \tag{3.8}$$

by reasoning just as above. (Here $M(z)$ plays the rôle that was played earlier by $\exp(-\beta\omega\log N)$.) This again corresponds to a non-interacting theory, but one with an arbitrary set of fundamental quanta of energy E_p, one for each prime p.

There is a myriad of other identities involving $\boldsymbol{\mu}(d)$. And one could, of course, construct physical systems and physical arguments to derive many of these results. But the above results are sufficient to give the flavor of the endeavor. However, it is worth observing that the particular physical system we have considered in this section seems to be a very natural one for representing number theoretic results. For example, consider the function $\Phi(\beta\omega, N)$ defined by the formula

$$\Phi(\beta\omega, N) = \mathrm{tr}\,[\delta(H-\omega N)(-1)^F e^{\beta H_b}] = \sum_d{}' \langle N, d|(-1)^F e^{\beta H_b}|N, d\rangle. \quad (3.9)$$

Evaluating (3.9) at the temperature $\beta^{-1} = \omega$, one obtains the function $\Phi(1, N)$, which turns out to be exactly the Euler totient function, the number of natural numbers less than N that are relatively prime to N. As another example, if one considers only the system defined by the bosonic Hamiltonian H_b, we can associate with any arithmetic function $\mathbf{h}(N)$ an operator $\tilde{h}$ such that $\langle N|\tilde{h}|N\rangle = \mathbf{h}(N)$. Then the function

$$\mathbf{h}'(N) = \frac{1}{\omega}\langle N|\tilde{h}H_b|N\rangle \quad (3.10)$$

is the Dirichlet derivative [3] of the arithmetic function $\mathbf{h}$. Thus we see that this quantum system seems to provide one of the most natural representations of number theoretic functions in physics.

4. The Möbius Inversion Formula

With the above experience, we can now study a question in number theory whose solution is the Möbius Inversion Formula. Most interestingly, we will see that once we have formulated the problem in physical terms, the answer immediately suggests itself, again in physical terms. This, we think, is an especially compelling feature of our construction.

We define an *arithmetic function* to be a function whose domain is the set of positive integers (i.e., natural numbers) and whose range is (a subset of) the complex numbers. Suppose we are given an arithmetic function $\mathbf{g}$. The question is, can we find an arithmetic function $\mathbf{h}$ such that

$$\mathbf{g}(n) = \sum_{m|n} \mathbf{h}(m), \quad (4.1)$$

where, as in the previous sections of this paper, "$\sum_{m|n}$" means to sum over all natural numbers m that divide n, and if so, how can we express $\mathbf{h}$ in terms of $\mathbf{g}$? We will see that there always is such an $\mathbf{h}$, and that the desired formula is easy to understand.

To analyze the problem, let us assume the function $\mathbf{h}$ exists, and let us use it to construct $\mathbf{g}$. So consider, to start with, a bosonic field theory which has creation operators $a_k^\dagger$, $k = 1, 2, 3, \ldots$ and with Hamiltonian H_a. Then the states of the theory are

$$\left(\prod_{j=1}^{r} (a_{i_j}^\dagger)^{\alpha_j}\right)|\mathrm{vacuum}\rangle \equiv |A\rangle, \quad (4.2)$$

where

$$A \equiv \prod_{j=1}^{r} (p_{i_j})^{\alpha_j}. \tag{4.3}$$

We can associate with any arithmetic function $\mathbf{h}$ an operator $\tilde{h}$; we define $\tilde{h}$ by specifying the expectation value of $\tilde{h}$ in any state, setting

$$\langle N|\tilde{h}|N\rangle = \mathbf{h}(N). \tag{4.4}$$

Now, what is the meaning of

$$\sum_{m|n} \mathbf{h}(m)? \tag{4.5}$$

This expression arises naturally if we enlarge the theory by adding an additional, entirely separate, bosonic piece H_b to the Hamiltonian. Thus we consider the Hamiltonian $H_a + H_b$, where H_a and H_b describe two completely separate, non-interacting systems. In addition to the creation operators $a_k^\dagger$, we now also have the creation operators $b_k^\dagger$ associated with the piece H_b of the field theory Hamiltonian. Then to specify a state in the theory we must keep track of both kinds of creation operators. A state is given by

$$\left(\prod_{j=1}^{r} (a_{i_j}^\dagger)^{\alpha_j} (b_{i_j}^\dagger)^{\beta_j}\right)|\text{vacuum}\rangle \equiv |A, B\rangle, \tag{4.6}$$

where

$$A \equiv (p_{i_1})^{\alpha_1} \cdot (p_{i_2})^{\alpha_2} \cdot \cdots \cdot (p_{i_r})^{\alpha_r}, \tag{4.7a}$$

$$B \equiv (p_{i_1})^{\beta_1} \cdot (p_{i_2})^{\beta_2} \cdot \cdots \cdot (p_{i_r})^{\beta_r}. \tag{4.7b}$$

(This notation is slightly different from the one we used in Sect. 2, where we derived the properties of the function $\boldsymbol{\mu}$. In that case, we considered the analogues not of A and B, but of AB and B.) The operator $\tilde{h}$ acting in the Hilbert space of the Hamiltonian H_a is now generalized to the operator $\tilde{h} \otimes \mathbf{1}$ acting on the Hilbert space of the combined Hamiltonian $H_a + H_b$; for convenience, however, we continue just to write the operator as $\tilde{h}$. (From context, it will always be clear on which space $\tilde{h}$ is acting, so there should be no confusion. Shortly, in fact, we will add a third sector to the theory, and we will still refer to this further generalized operator merely as $\tilde{h}$.)

We now can write down a representation of the function $\mathbf{g}$, namely

$$\mathbf{g}(N) = \sum_{A|N} \langle A|\tilde{h}|A\rangle = \sum_{AB=N} \langle A, B|\tilde{h}|A, B\rangle. \tag{4.8}$$

Thus $\mathbf{g}(N)$ is the sum of the expectation values of $\tilde{h}$ in all the states that are "N-times excited." So now, given $\mathbf{g}$, how do we recover $\mathbf{h}$? Note that by adding the piece H_b to the Hamiltonian, the sum over all "N-times excited" states now includes $\mathbf{h}(N)$, from the term involving $|N, 1\rangle$, as only one of many terms in the sum. Now we want to deconvolve this formula, that is, we want to remove the effects of these extra terms. To do this, clearly, we want to cancel out the effects of H_b. The answer is clear: as we know from BRST invariance [6, 7, 22] and supersymmetry [11, 23], the way to cancel the effects of a bosonic piece in the Hamiltonian H_b is to introduce as well a fermionic term H_f that is the superpartner of H_b.

So let us add to the Hamiltonian a piece H_f, associated with fermionic creation operators $f_k^\dagger$, $k = 1, 2, 3, \ldots$. Now the full Hamiltonian is $H_a + H_b + H_f$; but the expression $H_b + H_f$ is by itself supersymmetric. From the previous sections, we know that

$$\operatorname{tr}[(-1)^F e^{-\beta(H_b + H_f)}] = 1. \tag{4.9}$$

(Since $H_b + H_f$ is an arbitrary supersymmetric Hamiltonian, we choose a theory with a mass gap and a purely discrete spectrum to make the technicalities simple.) Then, let us consider a thermal expectation value of $\tilde{h}$ in the original theory. We see that

$$\begin{aligned}\langle \tilde{h} \rangle &= \operatorname{tr}[\tilde{h} e^{-\beta H_a}] \\ &= \operatorname{tr}[\tilde{h} e^{-\beta H_a}] \cdot \operatorname{tr}[(-1)^F e^{-\beta(H_b + H_f)}] \\ &= \operatorname{tr}[(-1)^F e^{-\beta(H_a + H_b + H_f)}],\end{aligned} \tag{4.10}$$

where the last identity follows because H_a does not interact with any part of the $H_b + H_f$ system. In principle, because this identity holds for all β, this formula gives us the desired answer, but it is clearer to approach things more concretely.

In our fully-enlarged theory, the states are

$$\left(\prod_{j=1}^{r} (a_{i_j}^\dagger)^{\alpha_j} (b_{i_j}^\dagger)^{\beta_j} (f_{i_j}^\dagger)^{\phi_j} \right) |\text{vacuum}\rangle \equiv |A, B, F\rangle, \tag{4.11}$$

where

$$A \equiv (p_{i_1})^{\alpha_1} \cdot (p_{i_2})^{\alpha_2} \cdot \cdots \cdot (p_{i_r})^{\alpha_r}, \tag{4.12a}$$

$$B \equiv (p_{i_1})^{\beta_1} \cdot (p_{i_2})^{\beta_2} \cdot \cdots \cdot (p_{i_r})^{\beta_r}, \tag{4.12b}$$

$$F \equiv (p_{i_1})^{\phi_1} \cdot (p_{i_2})^{\phi_2} \cdot \cdots \cdot (p_{i_r})^{\phi_r}. \tag{4.12c}$$

Note that the ϕ_j can only be 0 or 1, and so the allowed values of F are the squarefree positive integers, while A and B range over all the natural numbers.

The obvious expectation value to look at this point is

$$\sum_{ABF = N} (\langle A, B, F | \tilde{h} (-1)^F | A, B, F \rangle), \tag{4.13}$$

where the sum is over all "N-times excited" states in the larger theory. The idea is that the fermionic piece should cancel the bosonic piece, but of course the cancellation will only occur if we include a factor of $(-1)^F$. We can write the expectation value above in two ways, one involving the function **h** and one involving the function **g**. On the one hand,

$$\sum_{ABF = N} (\langle A, B, F | \tilde{h} (-1)^F | A, B, F \rangle) = \sum_{A | N} \left(\langle A | \tilde{h} | A \rangle \sum_{BF = N/A} \langle B, F | (-1)^F | B, F \rangle \right). \tag{4.14}$$

But,

$$\sum_{BF=N/A} \langle B,F|(-1)^F|B,F\rangle = \sum_{BF=N/A} (\boldsymbol{\mu}(F)) = \delta_{1,N/A}. \tag{4.15}$$

from (2.11) and (2.12). Thus

$$\sum_{ABF=N} (\langle A,B,F|\tilde{h}(-1)^F|A,B,F\rangle) = \sum_{A|N} \langle A|\tilde{h}|A\rangle\delta_{1,N/A} = \mathbf{h}(N). \tag{4.16}$$

On the other hand, grouping $H_a + H_b$ together and isolating H_f, we find

$$\begin{aligned}\sum_{ABF=N} (\langle A,B,F|\tilde{h}(-1)^F|A,B,F\rangle) &= \sum_{F|N}\left(\langle F|(-1)^F|F\rangle \sum_{AB=N/F} \langle A,B|\tilde{h}|A,B\rangle\right)\\ &= \sum_{F|N} \boldsymbol{\mu}(F)\mathbf{g}\left(\frac{N}{F}\right).\end{aligned} \tag{4.17}$$

The appearance of $\boldsymbol{\mu}$ comes from the identification of $\boldsymbol{\mu}$ with $(-1)^F$, while the appearance of $\mathbf{g}$ comes from (4.8).

Thus, combining the results (4.16) and (4.17), we find that

$$\mathbf{h}(N) = \sum_{d|N} \boldsymbol{\mu}(d)\mathbf{g}\left(\frac{N}{d}\right), \tag{4.18}$$

if, and only if,

$$\mathbf{g}(N) = \sum_{d|N} \mathbf{h}(d). \tag{4.19}$$

This result is known as the Möbius Inversion Formula. The result is very easy to understand now, as we have a supersymmetric interpretation of the formula. In constructing **g** from **h**, we are adding the effects of bosonic states; and so in obtaining **h** from **g**, we must essentially cancel the effects of these bosonic states, and the natural way to do this, from the modern point of view in theoretical physics, is to couple additional fermionic states to the theory to cancel the effects of the bosonic ones. Thus the appearance of the function $\boldsymbol{\mu}$, or $(-1)^F$, in the Möbius inversion formula is quite natural.

5. Conclusions

We have shown in this paper that supersymmetric quantum field theories (of the simplest kind) provide a natural context in which to understand certain results in number theory. This is illuminating for a variety of reasons, not the least of which is that it helps bring together the analyses of physics and number theory, and so sets up the possibility for further useful transfer of ideas. In addition, it is interesting to see that the Witten index (and the related bose-fermi cancellation in non-singlet representations of the supersymmetry algebra), already known to have topological and geometrical significance, can be used to understand a very different type of mathematical result. The Möbius inversion formula has to do with the arithmetic properties of the integers, with prime numbers and factorization. The key to making the connection between number theory and quantum systems developed in this paper lies in recognizing that the rôle played by creation operators to build up a Fock space is analogous to the rôle played by prime numbers to build up the natural numbers. From then on, the results follow naturally: fermion fields lead to the squarefree integers, $(-1)^F$ appears as $\boldsymbol{\mu}(d)$, and bose-fermi cancellation clearly can be used to unwrap the necessary expressions to yield the Möbius inversion formula. Still, that supersymmetric quantum field theory and the Witten index yield a result that is equivalent to the prime number theorem is truly surprising.

These results open up many possibilities. Further investigations of arithmetic functions in the physical context are warranted. One interesting class of functions which has a great deal of structure is the set of multiplicative functions, functions such that $\mathbf{f}(mn)=\mathbf{f}(m)\mathbf{f}(n)$ whenever m and n are relatively prime. One can develop a physical representation for this set of functions, which turn out to be systems which generalize the Einstein oscillator model of solids [19]. Obtaining further results on $\boldsymbol{\mu}$ from physical considerations also seems quite plausible. And, as we said above, understanding the infinite sums of $\boldsymbol{\mu}$, and in particular the asymptotic behavior of $\sum_{m=1}^{x} \boldsymbol{\mu}(m)$, would yield information regarding the Riemann hypothesis, whether all the non-trivial zeroes of the Riemann zeta function lie on the line $\mathrm{Re}(z)=\frac{1}{2}$.

In string theory, many important number theoretic functions appear, such as the classical partition function and various modular forms. It seems that any interesting number theoretic properties of these functions ought to have physical interpretations. An example of the type of question that we might want to understand in the context of string theory is, What is the physical property that can cause the Fourier coefficients of a modular form to be multiplicative, i.e., what is the physical interpretation of the Hecke operators [15]. Of course, we are as a whole still struggling to determine what the essential structure of string theory is, and how it is that string theory brings together so many different areas of mathematics—number theory, complex analysis, geometry, topology, and algebra. We hope the work in this paper helps us begin to understand in simpler contexts how number theory can be relevant in physical systems. We would not hope to understand fully the rôle of, say, geometry in string theory without experience in how geometry relates to physics in simpler contexts.

The $\boldsymbol{\mu}$ function has generalizations in other areas of mathematics. For example, there is a group theoretical Möbius inversion function, useful in the study of the subgroup content of groups. A physical interpretation of such a function seems quite likely.

The notion that mathematical problems have physical interpretations is growing in currency these days, though generally in the contexts of geometry and topology. We think this paper helps suggest that the parallels and isomorphisms between physics and math probably exist so as to contain all areas of math. Number theory seems much further removed, at first glance, from physics than does geometry or group theory. The fact that they are connected is both intriguing and suggestive, with promises of further discoveries to come.

Acknowledgements. I thank Mitchell Spector for introducing me to arithmetic function theory. This research was supported by a NATO Postdoctoral Fellowship from the U.S. National Science Foundation.

References

1. Akhoury, R., Comtet, A.: Anomalous Behavior of the Witten Index—Exactly Soluble Models. Nucl. Phys. **B246**, 253–278 (1984)
2. Alvarez-Gaumé, L.: Supersymmetry and the Atiyah–Singer Index Theorem. Commun. Math. Phys. **90**, 161–173 (1983)
3. Apostol, T. M.: An introduction to analytic number theory. Berlin, Heidelberg, New York: Springer 1976
4. Apostol, T. M.: Modular functions and dirichlet series in number theory. Berlin, Heidelberg, New York: Springer 1976
5. Ayoub, R. G.: An introduction to the analytic theory of numbers. Mathematical Surveys, No. 10. Providence, R.I.: American Mathematical Society 1963
6. Becchi, C., Rouet, A., Stora, R.: The Abelian Higgs–Kibble model, unitarity of the S-operator. Phys. Lett. **52B**, 344–346 (1974)
7. Becchi, C., Rouet, A., Stora, R.: Renormalization of gauge theories. Ann. Phys. **98**, 297–321 (1976)
8. Candelas, P., Horowitz, G., Strominger, A., Witten, E.: Vacuum configurations for superstrings. Nucl. Phys. **B258**, 46–74 (1985)
9. Cecotti, S., Girardello, L.: Functional measure, topology, and dynamical supersymmetry breaking. Phys. Lett. **110B**, 39–43 (1982)
10. Enderton, H. B.: A mathematical introduction to logic. New York: Academic Press 1972
11. Gates, S. J., Jr., Grisaru, M., Rocek, M., Siegel, W.: Superspace. Reading MA.: Bemjamin/Cummings 1983
12. Georgi, H. M.: Lie algebras and particle Physics. Reading, MA.: Benjamin/Cummings 1982
13. Green, M. B., Schwarz, J. H., Witten, E.: Superstring theory, Vol. I and II. Cambridge, New York: Cambridge University Press 1987
14. Hagedorn, R.: Hadronic matter near the boiling point. Nuovo Cimento **56A**, 1027–1057 (1968)
15. Moore, G.: Atkin–Lehner symmetry. Nucl. Phys. **B293**, 139–188 (1987)
16. Nicolai, H.: Supersymmetry and functional integration measures. Nucl. Phys. **B176**, 419–428 (1980)
17. Niven, I., Zuckerman, H. S.: An introduction to the theory of numbers, 3rd (ed). New York: John Wiley 1972
18. Serre, J.-P.: A course in arithmetic. Berlin, Heidelberg, New York: Springer 1973
19. Spector, D.: Multiplicative functions, dirichlet convolution, and quantum systems. Phys. Lett. A. (in press)
20. Tinkham, M.: Group theory and quantum mechanics. New York: McGram–Hill 1975
21. Titchmarsh, E. C.: The theory of the Riemann zeta function. Oxford: Clarendon Press 1951
22. Tyupin, I. V.: Gauge invariance in field theory and in statistical physics in the operator formulation. Lebedev preprint FIAN No. 39 (1975)
23. Wess, J., Bagger, J.: Supersymmetry and supergravity. Princeton, NJ: Princeton University Press 1983
24. Witten, E.: Constraints on supersymmetry breaking. Nucl. Phys. **B202**, 253–316 (1982)
25. Witten, E.: Topological quantum field theory. Commun. Math. Phys. **117**, 353–386 (1988)
26. Witten, E.: Quantum field theory and the Jones polynomial. Commun. Math. Phys.

Communicated by L. Alvarez-Gaumé

Received February 20, 1989