ON FINITE SUMS OF RECIPROCALS OF DISTINCT *n*TH POWERS

R. L. GRAHAM

Introduction. It has long been known that every positive rational number can be represented as a finite sum of reciprocals of distinct positive integers (the first proof having been given by Leonardo Pisano [6] in 1202). It is the purpose of this paper to characterize (cf. Theorem 4) those rational numbers which can be written as finite sums of reciprocals of distinct nth *powers* of integers, where n is an arbitrary (fixed) positive integer and "finite sum" denotes a sum with a finite number of summands. It will follow, for example, that p/q is the finite sum of reciprocals of distinct squares[1] if and only if

$$\frac{p}{q} \in \left[0, \frac{\pi^2}{6} - 1\right) \cup \left[1, \frac{\pi^2}{6}\right).$$

Our starting point will be the following result:

THEOREM A. *Let n be a positive integer and let H^n denote the sequence $(1^{-n}, 2^{-n}, 3^{-n}, \cdots)$. Then the rational number p/q is the finite sum of distinct terms taken from H^n if and only if for all $\varepsilon > 0$, there is a finite sum s of distinct terms taken from H^n such that $0 \leqq s - p/q < \varepsilon$.*

Theorem A is an immediate consequence of a result of the author [2, Theorem 4] together with the fact that every sufficiently large integer is the sum of distinct nth powers of positive integers (cf., [8], [7] or [3]).

The main results. We begin with several definitions. Let $S = (s_1, s_2, \cdots)$ denote a (possibly finite) sequence of real numbers.

DEFINITION 1. $P(S)$ is defined to be the set of all sums of the form $\sum_{k=1}^{\infty} \varepsilon_k s_k$ where $\varepsilon_k = 0$ or 1 and all but a finite number of the ε_k are 0.

DEFINITION 2. $Ac(S)$ is defined to be the set of all real numbers x such that for all $\varepsilon > 0$, there is an $s \in P(S)$ such that $0 \leqq s - x < \varepsilon$. Note that in this terminology Theorem A becomes:

$$(1) \qquad P(H^n) = Ac(H^n) \cap Q$$

Received May 13, 1963.

[1] This result has also been obtained by P. Erdös (not published).

where Q denotes the set of rational numbers.

DEFINITION 3. A term s_n of S is said to be *smoothly replaceable in* S (abbreviated s.r. *in* S) if $s_n \leqq \sum_{k=1}^{\infty} s_{n+k}$.

THEOREM 1. *Let* $S = (s_1, s_2, \cdots)$ *be a sequence of real numbers such that:*

1. $s_n \downarrow 0$.
2. *There exists an integer* r *such that* $n \geqq r$ *implies that* s_n *is smoothly replaceable in* S.

Then

$$Ac(S) = \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$$

where $P_{r-1} = P((s_1, \cdots, s_{r-1}))$ (*note that* $P_0 = \{0\}$) *and* $\sigma = \sum_{k=r}^{\infty} s_k$(*where possibly* σ *is infinite*).

Proof. Let $x \in \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$ and assume that $x \notin Ac(S)$. Then $x \in [\pi, \pi + \sigma)$ for some $\pi \in P_{r-1}$. A sum of the form $\pi + \sum_{t=1}^{k} s_{i_t}$ where $r \leqq i_1 < i_2 < \cdots < i_k$ will be called "minimal" if

$$\pi + \sum_{t=1}^{k-1} s_{i_t} < x < \pi + \sum_{t=1}^{k} s_{i_t} \tag{2}$$

(where a sum of the form $\sum_{t=a}^{b}$ is taken to be 0 for $b < a$). Note that since $x \notin Ac(S) \supset P(S)$ then we never get equality in (2). Let M denote the set of minimal sums. Then M must contain infinitely many elements. For suppose M is a finite set. Let m denote the largest index of any s_j which is used in any element of M and let $p = \pi + \sum_{k=1}^{n} s_{j_k} + s_m$ be an element of M which uses s_m (where $r \leqq j_1 < j_2 < \cdots < j_n < m$ and possibly n is zero). Thus we have

$$\pi + \sum_{k=1}^{n} s_{j_k} < x < \pi + \sum_{k=1}^{n} s_{j_k} + \sum_{t=1}^{\infty} s_{m+t}$$

since s_m is s.r. in S. Therefore there is a *least* $d \geqq 1$ such that $x < p' = \pi + \sum_{k=1}^{n} s_{j_k} + \sum_{t=1}^{d} s_{m+t}$. Hence p' is a "minimal" sum which uses s_{m+d} and $m + d > m$. This is a contradiction to the definition of m and consequently M must be infinite. Now, let $\delta = \inf\{p - x: p \in M\}$. Since $x \notin Ac(S)$ then $\delta > 0$. There exist $p_1, p_2, \cdots \in M$ such that $p_n - x < \delta + \delta/2^n$. Since $s_n \downarrow 0$ then there exists c such that $n \geqq c$ implies that $s_n < \delta/2$. Also, there exists w such that $n \geqq w$ implies that p_n uses an s_k for some $k \geqq c$ (since only a finite number of p_j can be formed from the s_k with $k < c$). Therefore we can write $p_w = \pi + \sum_{j=1}^{n} s_{k_j}$ where $k_n \geqq c$. Hence

$$p_w - s_{k_n} - x > p_w - \frac{\delta}{2} - x \geqq \delta - \frac{\delta}{2} = \frac{\delta}{2} > 0$$

which is a contradiction to the assumption that p_w is "minimal." Thus, we must have $x \in Ac(S)$ and consequently

$$\bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma) \subset Ac(S) . \tag{3}$$

To show inclusion in the other direction let $x \in Ac(S)$ and suppose that $x \notin \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$. Thus, either $x < 0$, $x \geqq \sum_{k=1}^{\infty} s_k$, or there exist π and π' in P_{r-1} such that $\pi + \sigma \leqq x < \pi'$ where no element of P_{r-1} is contained in the interval $[\pi + \sigma, \pi')$. Since the first two possibilities imply that $x \notin Ac(S)$ (contradicting the hypothesis) then we may assume that the third possibility holds. Therefore there exists $\delta > 0$ such that

$$x \leqq \pi' - \delta . \tag{4}$$

Let p be any element of $P(S)$. Then $p = \sum_{t=1}^{m} s_{i_t} + \sum_{u=1}^{n} s_{j_u}$ for some m and n where

$$1 \leqq i_1 < i_2 < \cdots < i_m \leqq r - 1 < j_1 < j_2 < \cdots < j_n .$$

Thus for $\pi^* = \sum_{t=1}^{m} s_{i_t}$ we have $p \in [\pi^*, \pi^* + \sigma)$. Consequently any element p of $P(S)$ must fall into an interval $[\pi^*, \pi^* + \sigma)$ for some $\pi^* \in P_{r-1}$ and therefore, if p exceeds x then it must exceed x by at least δ (since $p \notin [\pi + \sigma, \pi')$ and thus by (4) $p > x \in [\pi + \sigma, \pi')$ implies $p \geqq \pi' \geqq x + \delta$). This contradicts the hypothesis that $x \in Ac(S)$ and hence we conclude that $Ac(S) \subset \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$. Thus, by (3) we have $Ac(S) = \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$ and the theorem is proved.

THEOREM 2. *Let $S = (s_1, s_2, \cdots)$ be a sequence of real numbers such that:*

1. $s_n \downarrow 0$.
2. *There exists an integer r such that $n < r$ implies that s_n is not* s.r. *in S while $n \geqq r$ implies that s_n is* s.r. *in S.*

Then $Ac(S)$ is the disjoint union of exactly 2^{r-1} half-open intervals each of length $\sum_{k=r}^{\infty} s_k$.

Proof. By Theorem 1 we have $Ac(S) = \bigcup_{\pi \in P_{r-1}} [\pi, \pi + \sigma)$ where $\sigma = \sum_{k=r}^{\infty} s_k$ and $P_{r-1} = P((s_1, \cdots, s_{r-1}))$. Let $\pi = \sum_{k=1}^{u} s_{i_k}$ and $\pi' = \sum_{k=1}^{v} s_{j_k}$ be any two formally distinct sums of the s_n where $1 \leqq i_1 < \cdots < i_u \leqq r - 1$ and $1 \leqq j_1 < \cdots < j_v \leqq r - 1$ and we can assume without loss of generality that $\pi \geqq \pi'$. Then either there is a *least* $m \geqq 1$ such that $i_m \neq j_m$ or we have $i_k = j_k$ for $k = 1, 2, \cdots, v$ and

$u > v$. In the first case we have

$$\begin{aligned}\pi &= \sum_{k=1}^{u} s_{i_k} = \sum_{k=1}^{m-1} s_{j_k} + \sum_{k=m}^{u} s_{i_k} \\ &> \sum_{k=1}^{m-1} s_{j_k} + \sum_{k=1}^{\infty} s_{i_m+k} \text{ (since } s_{i_m} \text{ is not s.r. in } S) \\ &\geqq \pi' + \sigma \text{ (since } j_m \geqq i_m + 1) .\end{aligned}$$

In the second case we have

$$\begin{aligned}\pi &= \sum_{k=1}^{u} s_{i_k} = \sum_{k=1}^{v} s_{j_k} + \sum_{k=v+1}^{u} s_{i_k} \\ &> \sum_{k=1}^{v} s_{j_k} + \sum_{k=1}^{\infty} s_{i_{v+1}+k} \text{ (since } s_{i_{v+1}} \text{ is not s.r. in } S) \\ &\geqq \pi' + \sigma \text{ (since } i_{v+1}+1 \leqq i_u + 1 \leqq r) .\end{aligned}$$

Thus, in either case we see that $\pi > \pi' + \sigma$. Consequently, any two formally distinct sums in P_{r-1} are separated by a distance of more than σ and hence, each element π of P_{r-1} gives rise to a half-open interval $[\pi, \pi + \sigma)$ which is disjoint from any other interval $[\pi', \pi' + \sigma)$ for $\pi \neq \pi' \in P_{r-1}$. Therefore $Ac(S) = \bigcup_{\pi \in P_{r-1}}[\pi, \pi + \sigma)$ is the disjoint union of exactly 2^{r-1} half-open intervals $[\pi, \pi + \sigma)$, $\pi \in P_{r-1}$, (since there are exactly 2^{r-1} formally distinct sums of the form $\sum_{k=1}^{r-1} \varepsilon_k s_k$, $\varepsilon_k =$ 0 or 1) where each interval is of length σ. This proves the theorem.

We now need three additional lemmas in order to prove the main theorems.

LEMMA 1. *Let $S = (s_1, s_2, \cdots)$ be a sequence of nonnegative real numbers and suppose that there exists an m such that $n \geqq m$ implies that $s_n \leqq 2s_{n+1}$. Then $n \geqq m$ implies that s_n is* s.r. *in S* (i.e., $s_n \leqq \sum_{k=1}^{\infty} s_{n+k}$).

Proof. If $\sum_{k=1}^{\infty} s_k = \infty$ then the lemma is immediate. Assume that $\sum_{k=1}^{\infty} s_k < \infty$. Then

$$n \geqq m \Longrightarrow s_{n+k} \geqq \frac{1}{2} s_{n+k-1} , \qquad k = 1, 2, 3, \cdots$$

$$\Longrightarrow \sum_{k=1}^{\infty} s_{n+k} \geqq \frac{1}{2} \sum_{k=1}^{\infty} s_{n+k-1} = \frac{1}{2} s_n + \frac{1}{2} \sum_{k=1}^{\infty} s_{n+k} .$$

Therefore, $s_n \leqq \sum_{k=1}^{\infty} s_{n+k}$, i.e., s_n is s.r. in S.

LEMMA 2. *Suppose that $k \leqq (2^{1/n} - 1)^{-1}$ and k^{-n} is* s.r. *in H^n (where H^n was defined to be the sequence $(1^{-n}, 2^{-n}, \cdots)$). Then $(k + 1)^{-n}$ is also* s.r. *in H^n.*

Proof.

$$k \leqq (2^{1/n} - 1)^{-1} \Longrightarrow \frac{1}{k} \leqq 2^{1/n} - 1$$

$$(5) \qquad \Longrightarrow \left(1 + \frac{1}{k}\right)^n \geqq 2$$

$$\Longrightarrow k^{-n} \geqq 2(k+1)^{-n} .$$

Since by hypothesis, $\sum_{j=k+1}^{\infty} j^{-n} \geqq k^{-n}$, then by (5)

$$\sum_{j=k+2}^{\infty} j^{-n} \geqq k^{-n} - (k+1)^{-n} \geqq 2(k+1)^{-n} - (k+1)^{-n} = (k+1)^{-n} .$$

Hence, $(k+1)^{-n}$ is s.r. in H^n and the lemma is proved.

LEMMA 3. Suppose that $k \geqq (2^{1/n} - 1)^{-1}$. Then k^{-n} is s.r. in H_n.

Proof.

$$r \geqq k \Longrightarrow r \geqq (2^{1/n} - 1)^{-1}$$

$$\Longrightarrow \frac{1}{r} \leqq 2^{1/n} - 1$$

$$\Longrightarrow \left(1 + \frac{1}{r}\right)^n \leqq 2$$

$$\Longrightarrow r^{-n} \leqq 2(r+1)^{-n} .$$

Therefore, by Lemma 1, r^{-n} is s.r. in H^n.

THEOREM 3. *Let t_n denote the largest integer k such that k^{-n} is not* s.r. *in H^n and let P denote $P((1^{-n}, 2^{-n}, \cdots, t_n^{-n}))$. Then*

$$Ac(H^n) = \bigcup_{\pi \in P} [\pi, \pi + \sum_{k=1}^{\infty} (t_n + k)^{-n})$$

is the disjoint union of exactly 2^{t_n} intervals. Moreover, $t_n < (2^{1/n} - 1)^{-1}$ and $t_n \sim n/ln\, 2$ (where ln 2 denotes $\log_e 2$).

Proof. With the exception of $t_n \sim n/ln\, 2$, the theorem follows directly from the preceding results. The following argument, due to L. Shepp, shows that $t_n \sim n/ln\, 2$.

Consider the function $f_n(x)$ defined by

$$(6) \qquad f_n(x) = x^n\left(\sum_{k=1}^{\infty} \frac{1}{(x+k)^n} - \frac{1}{x^n}\right)$$

for $n = 2, 3, \cdots$ and $x > 0$. Since

$$f_n(x) = \sum_{k=1}^{\infty} \left(1 + \frac{k}{x}\right)^{-n} - 1$$

then $f_n(x) < 0$ for sufficiently small $x > 0$, $f_n(x) > 0$ for sufficiently large x, and $f_n(x)$ is continuous and monotone increasing for $x > 0$. Hence the equation $f_n(x) = 0$ has a unique positive root x_n and from the definition of t_n it follows by (6) that $0 < x_n - t_n \leqq 1$. Thus, to show that $t_n \sim n/ln\,2$, it suffices to show that $x_n \sim n/ln\,2$. Now it is easily shown (cf., [4], p. 13) that for $a > 0$, $(1 + \alpha/n)^{-n}$ is a decreasing function of n. Thus, $f_n(\alpha n)$ is a decreasing function of n and since $f_2(2\alpha) < \infty$ for $\alpha > 0$ then

$$\begin{aligned}\lim_{n\to\infty} f_n(\alpha n) &= \lim_{n\to\infty}\sum_{k=1}^{\infty}\left(1+\frac{k}{\alpha n}\right)^{-n} - 1\\ &= \sum_{k=1}^{\infty}\lim_{n\to\infty}\left(1+\frac{k}{\alpha n}\right)^{-n} - 1\\ &= -1 + \sum_{k=1}^{\infty} e^{-k/\alpha} = (e^{1/\alpha}-1)^{-1} - 1\end{aligned}$$

since the monotone convergence theorem (cf., [5]) allows us to interchange the sum and limit. Suppose now that for some $\varepsilon > 0$, there exist $n_1 < n_2 < \cdots$ such that $x_{n_i} > n_i(1/ln\,2 + \varepsilon)$. Then

$$\begin{aligned}0 = \lim_{i\to\infty} f_{n_i}(x_{n_i}) &\geqq \lim_{i\to\infty} f_{n_i}\left(n_i\left(\frac{1}{ln\,2}+\varepsilon\right)\right)\\ &= (e^{(1/ln\,2+\varepsilon)^{-1}} - 1)^{-1} - 1\\ &= (2^{1/(1+\varepsilon\, ln\,2)} - 1)^{-1} - 1 > 0\end{aligned}$$

which is a contradiction. Similarly, if for some ε, $0 < \varepsilon < 1/ln\,2$, there exist $n_1 < n_2 < \cdots$ such that

$$x_{n_i} < n_i\left(\frac{1}{ln\,2} - \varepsilon\right),$$

then

$$\begin{aligned}0 = \lim_{i\to\infty} f_{n_i}(x_{n_i}) &\leqq \lim_{i\to\infty} f_{n_i}\left(n_i\left(\frac{1}{ln\,2}-\varepsilon\right)\right)\\ &= (e^{(1/ln\,2-\varepsilon)^{-1}} - 1)^{-1} - 1\\ &= (2^{1/(1-\varepsilon\, ln\,2)} - 1)^{-1} - 1 < 0\end{aligned}$$

which is again impossible. Hence we have shown that for all $\varepsilon > 0$, there exists an n_0 such that $n > n_0$ implies that

$$n\left(\frac{1}{ln\,2} - \varepsilon\right) \leqq x_n \leqq n\left(\frac{1}{ln\,2} + \varepsilon\right)$$

or equivalently

$$-\varepsilon \leqq \frac{x_n}{n} - \frac{1}{ln\,2} \leqq \varepsilon\,.$$

Therefore, $\lim_{n\to\infty} x_n/n = 1/ln\,2$ and the theorem is proved.[2]

The following table gives the values of t_n for some small values of n.

n	t_n	$[(2^{1/n}-1)^{-1}]$
1	0	1
2	1	2
3	2	3
4	4	5
5	5	6
10	12	13
100	?	143
1000	?	1442

We may now combine Theorem 3 and Theorem A (cf. Eq. (1)) and express the result in ordinary terminology to give:

THEOREM 4. *Let n be a positive integer, let t_n be the largest integer k such that $k^{-n} > \sum_{j=1}^{\infty}(k+j)^{-n}$ and let P denote the set $\{\sum_{j=1}^{t_n}\varepsilon_j j^{-n}\colon \varepsilon_j = 0$ or $1\}$. Then the rational number p/q can be written as a finite sum of reciprocals of distinct* nth *powers of integers if and only if*

$$\frac{p}{q} \in \bigcup_{\pi\in P}\left[\pi, \pi + \sum_{k=1}^{\infty}(t_n+k)^{-n}\right) .$$

COROLLARY 1. *p/q can expressed as the finite sum of reciprocals of distinct squares if and only if*

$$\frac{p}{q} \in \left[0, \frac{\pi^2}{6} - 1\right) \cup \left[1, \frac{\pi^2}{6}\right) .$$

COROLLARY 2. *p/q can be expressed as the finite sum of reciprocals of distinct cubes if and only if*

$$\frac{p}{q} \in \left[0, \zeta(3) - \frac{9}{8}\right) \cup \left[\frac{1}{8}, \zeta(3) - 1\right) \cup \left[1, \zeta(3) - \frac{1}{8}\right) \cup \left[\frac{9}{8}, \zeta(3)\right)$$

where $\zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.2020569\cdots$

REMARKS. In theory it should be possible to calculate directly from the relevant theorems (cf., [2], [3]) an explicit bound for the number of terms of H^n needed to represent p/q as an element of $P(H^n)$. However, since the theorems were not designed to minimize such a bound, but rather merely to show its existence, then understandably, this calculated bound would probably be many orders of

[2] In fact, it can be shown that x_n has the expansion $n/\ln 2 - 1/2 + c_1 n^{-1} + \cdots + c_k n^{-k} + 0(n^{-k-1})$ for any k.

magnitude too large. Erdös and Stein [1] and, independently, van Albada and van Lint [9] have shown that if $f(n)$ denotes the least number of terms of $H^1 = (1^{-1}, 2^{-1}, \cdots)$ needed to represent the integer n as an element of $P(H^1)$ then $f(n) \sim e^{n-\gamma}$ where γ is Euler's constant.

It should be pointed out that a more general form of Theorem A may be derived from [2] which can be used to prove results of the following type:

COROLLARY A. *The rational p/q with $(p, q) = 1$ can be expressed as a finite sum of reciprocals of distinct odd squares if and only if q is odd and $p/q \in [0, (\pi^2/8) - 1) \cup [1, \pi^2/8)$.*

COROLLARY B. *The rational p/q with $(p, q) = 1$ can be expressed as a finite sum of reciprocals of distinct squares which are congruent to 4 modulo 5 if and only if $(q, 5) = 1$ and*

$$\frac{p}{q} \in \left[0, \alpha - \frac{13}{36}\right) \cup \left[\frac{1}{9}, \alpha - \frac{1}{4}\right) \cup \left[\frac{1}{4}, \alpha - \frac{1}{9}\right) \cup \left[\frac{13}{36}, \alpha\right)$$

where $\alpha = 2(5 - \sqrt{5})\pi^2/125 = \sum_{k=0}^{\infty}((5k+2)^{-2} + (5k+3)^{-2}) = 0.43648\cdots$

It is not difficult to obtain representations of specific rationals as elements of $P(H^n)$ (for small n), e.g.,

$$\frac{1}{2} = 2^{-2} + 3^{-2} + 4^{-2} + 5^{-2} + 6^{-2} + 15^{-2} + 18^{-2} + 36^{-2} + 60^{-2} + 180^{-2}\,,$$

$$\frac{1}{3} = 2^{-2} + 4^{-2} + 10^{-2} + 12^{-2} + 20^{-2} + 30^{-2} + 60^{-2}\,,$$

$$\frac{5}{37} = 2^{-3} + 5^{-3} + 10^{-3} + 15^{-3} + 16^{-3} + 74^{-3} + 111^{-3} + 185^{-3} + 240^{-3} + 296^{-2} + 444^{-3} + 1480^{-3}\,, \text{ etc.!}$$

REFERENCES

1. P. Erdös and S. Stein, *Sums of distinct unit fractions*, Proc. Amer. Math. Soc., **14** (1963), 126-131.
2. R. Graham, *On finite sums of unit fractions*, Proc. London Math. Soc., (to appear).
3. ———, *Complete sequences of polynomial values*, Duke Math. Jour. (to appear).
4. P. P. Korovkin, *Inequalities*, Random House, New York (1961).
5. M. Loève, *Probability Theory*, Van Nostrand, Princeton, (1960).
6. Leonardo Pisano, *Scritti*, vol. 1, B. Boncompagni, Rome (1857) 77-83.
7. K. F. Roth and G. Szekeres, *Some asymptotic formulae in the theory of partitions*, Quart. Jour. of Math., **5** (1954), 241-259.
8. R. Sprague, *Über Zerlegungen in n-te Potenzen mit lauter verschiedene Grundzahlen*, Math. Zeit., **51** (1947-48), 466-468.
9. P. J. van Albada and J. H. van Lint, *Reciprocal bases for the integers*, Amer. Math. Monthly **70** (1963), 170-174.

BELL TELEPHONE LABORATORIES, INC.