On the number of hypercubic bipartitions of an integer

Geir Agnarsson *

Abstract

We revisit a well-known divide-and-conquer maximin recurrence $f(n) = \max(\min(n_1, n_2) + f(n_1) + f(n_2))$ where the maximum is taken over all proper bipartitions $n = n_1 + n_2$, and we present a new characterization of the pairs (n_1, n_2) summing to n that yield the maximum $f(n) = \min(n_1, n_2) + f(n_1) + f(n_2)$. This new characterization allows us, for a given $n \in \mathbb{N}$, to determine the number h(n) of these bipartitions that yield the said maximum f(n). We present recursive formulae for h(n), a generating function h(x), and an explicit formula for h(n) in terms of a special representation of n.

2000 MSC: 05A15, 05C35.

Keywords: bipartition, rectangular grid, hypercube, induced subgraphs, divide-and-conquer, divide-and-conquer maximin recurrence.

1 Introduction

The purpose of this article is to further contribute to the study of the maximum number of edges of an induced subgraph on n vertices of the hypercube Q_k . In order to do that, we revisit a wellknown divide-and-conquer maximin recurrence and recap some of its properties. As stated later in the section, it is easy to see that, more generally, the solution to this mentioned recurrence equals the maximum number of edges an induced subgraph on n vertices in a rectangular grid \mathbb{Z}^k can have. These considerations were in part initially inspired by the heuristic integer sequence $0, 1, 2, 4, 5, 7, 9, 12, 13, 15, 17, 20, \ldots$ [1, A007818], describing the maximal number of edges joining $n = 1, 2, 3, \ldots$ vertices in the cubic rectangular grid \mathbb{Z}^3 , for which no general formula nor procedure to compute it is given. – First we set forth our basic terminology and definitions.

Notation and terminology The set of integers will be denoted by \mathbb{Z} , the set of natural numbers $\{1, 2, 3, \ldots\}$ by \mathbb{N} , and the set of non-negative integers $\{0, 1, 2, 3, \ldots\}$ by \mathbb{N}_0 . The base-two logarithm of a real x will be denoted by $\lg x$. Unless otherwise stated, all graphs in this article will be finite, simple and undirected. For a graph G, its set of vertices will be denoted by V(G) and its set of edges by E(G). Clearly $E(G) \subseteq \binom{V(G)}{2}$ the set of all 2-element subsets of V(G). We will denote an edge with endvertices u and v by uv instead of the actual 2-set $\{u, v\}$. The order of G is |V(G)| and the size of G is |E(G)|. By an induced subgraph H of G we mean a subgraph H such that $V(H) \subseteq V(G)$ in the usual set theoretic sense, and such that if $u, v \in V(H)$ and $uv \in E(G)$, then $uv \in E(H)$. If $U \subseteq V(G)$ then the subgraph of G induced by V will be denoted by G[U].

^{*}Department of Mathematical Sciences, George Mason University, MS 3F2, 4400 University Drive, Fairfax, VA 22030, geir@math.gmu.edu

For $k \in \mathbb{N}$ a rectangular grid \mathbb{Z}^k in our context is a infinite graph with the point set \mathbb{Z}^k as its vertices and where two points $\tilde{x} = (x_1, \ldots, x_k)$ and $\tilde{y} = (y_1, \ldots, y_k)$ are connected by an edge iff the Manhattan distance $d(\tilde{x}, \tilde{y}) = \sum_{i=1}^{k} |x_i - y_i| = 1$. So, two points are connected iff they only differ in one coordinate, in which they differ by ± 1 . This Manhattan distance measure is the metric corresponding to the 1-norm $\|\tilde{x}\|_1 = \sum_{i=1}^{k} |x_i|$ in the k-dimensional Euclidean space \mathbb{R}^k . The hypercube Q_k is then the subgraph of the grid \mathbb{Z}^k induced by the 2^k points $\{0,1\}^k$. The vertices of the hypercube Q_k are more commonly viewed as binary strings of length k instead of actually points in the k-dimensional Euclidean space. In that case the Manhattan distance is called the called the Hamming distance. We will not make a specific distinction between these two slightly different presentations of the hypercube Q_k .

Assume we have as set of n distinct vertices $U = \{u_1, \ldots, u_n\}$ in a rectangular grid and consider the induced graph G[U] with the maximum number of edges among all induced subgraphs on n vertices of the rectangular grid. We will assume its dimension k to be large enough so our considerations will not be hindered by its value in any way. Each vertex u_i is represented by a point $\tilde{x}_i = (x_{i1}, \ldots, x_{ik})$ in \mathbb{Z}^k , so we may assume that the j-th coordinates $x_{1j}, x_{2j}, \ldots, x_{nj}$ are not all identical, since otherwise no induced edge in G[U] will be lost by projection $\pi_j : \mathbb{Z}^k \to \mathbb{Z}^{k-1}$ where the j-th coordinate has been removed. In other words, we may assume that $|\pi_j(U)| \ge 2$ for each coordinate j. In particular, there is a proper partition $U = U_1 \cup U_2$ where the first coordinate of each vertex in U_1 is less than the first coordinate of each vertex in U_2 . Let $|U_1| = n_1$ and $|U_2| = n_2$, so $n = n_1 + n_2$ where $n_1, n_2 \ge 1$. Assuming that G[U] has the maximum number g(n) of edges such an induced graph in the grid can have, we can have at most $\min(n_1, n_2)$ edges in G[U]parallel to the first coordinate axis. Since each edge in G[U] between a vertex of U_1 and a vertex of U_2 must be parallel to the first coordinate axis, all other edges in G[U] are in the disjoint union $E(G[U_1]) \cup E(G[U_2])$, which ideally are maximally connected. Trivially we have g(1) = 0 and so we have the following.

Observation 1.1 For $n \in \mathbb{N}$ let g(n) denote the maximum number of edges an induced subgraph on n vertices of \mathbb{Z}^k can have. Then g(1) = 0 and

$$g(n) \le \max_{\substack{n_1+n_2=n\\n_1,n_2\ge 1}} \left(\min(n_1,n_2) + g(n_1) + g(n_2)\right).$$

REMARK: It is apriori not clear that we have equality in Observation 1.1, since by insisting that we have the maximum number of edges the first coordinate in the k-dimensional grid allows, this restricts the structure of both the "upper" and the "lower" induced subgraphs with on fewer vertices.

A recap of well-known results We next recap some known relevant properties about the function f on the nonnegative integers given by the following divide-and-conquer recursion.

$$f(n) \begin{cases} 0 \text{ if } n = 1, \\ \lfloor n/2 \rfloor + f(\lfloor n/2 \rfloor) + f(\lceil n/2 \rceil) \text{ if } n > 1. \end{cases}$$
(1)

This function f and its number sequence $(f(n))_{n=0}^{\infty} = (0, 1, 2, 4, 5, 7, 9, 12, 13, 15, 17, 20, ...)$ is a well-known sequence as given in [2, A000788] where it is presented by a slightly different recursion. In fact, the behaviour of f(n) is interesting also from analytic point of view as $f(n) = \frac{n}{2} \lg n + O(n)$. The asymptotic behavior of $\frac{n}{2} \lg n - f(n)$ was first studied in detail in [13] and it has fractal-like

shape. It tends to towards the Blancmange function [3], a continuous function which is nowhere differentiable [4] on every interval $[2^i, 2^{i+1}]$ between two consecutive powers of 2. The Blancmange-like graph of $f(n) - \frac{n}{2} \lg n$ (a negative function) also appears in [6, Fig. 1, p. 256].

If s(n) denotes the sum of the digits of n when expressed as a binary number (or just the number of 1s appearing in the binary expression of n), then clearly s(n) = s(n-1) + 1 when n is odd, and s(n) = s(n/2) when n is even, and therefore

$$s(n) = \begin{cases} s((n-1)/2) + 1 & \text{if } n \text{ is odd }, \\ s(n/2) & \text{if } n \text{ is even }. \end{cases}$$
(2)

Also, when we express all the *n* integers 0, 1, ..., n-1 as binary numbers, $\lfloor n/2 \rfloor$ of them are odd and $\lceil n/2 \rceil$ even. From this and (2) it is evident that

$$\sum_{i=0}^{n-1} s(i) = \sum_{l=0}^{\lfloor n/2 \rfloor - 1} s(2l+1) + \sum_{l=0}^{\lceil n/2 \rceil - 1} s(2l) = \lfloor n/2 \rfloor + \sum_{l=0}^{\lfloor n/2 \rfloor - 1} s(l) + \sum_{l=0}^{\lceil n/2 \rceil - 1} s(l),$$

which is same recursion that f satisfies. Hence, we have the following as stated in [2, A000788] and [6].

Observation 1.2 For $n \in \mathbb{N}$ we have $f(n) = \sum_{i=0}^{n-1} s(i)$.

For $n \in \mathbb{N}$ the number of digits in the binary expression of n-1 is $k = \lceil \lg n \rceil$. For each $i \in \{0, 1, \ldots, n-1\}$ there is a corresponding binary point $\tilde{\beta}_k(i) \in \{0, 1\}^k$ from the binary expression of i where the last digit of i is the k-th coordinate, the next to last digit is the (k-1)-th coordinate and so forth.

Proposition 1.3 For $n \in \mathbb{N}$ and $k = \lceil \lg n \rceil$ the *n* points $\{\tilde{\beta}_k(0), \tilde{\beta}_k(1), \dots, \tilde{\beta}_k(n-1)\}$ induce a subgraph in Q_k with f(n) edges.

Proof. For each $i \in \{1, \ldots, n-1\}$ we note that $\hat{\beta}_k(i)$ is connected by and edge to exactly s(i) previous points $\tilde{\beta}_k(0), \ldots, \tilde{\beta}_k(i-1)$, namely, those s(i) points obtained from $\tilde{\beta}_k(i)$ by replacing each of the 1s by a 0. By Observation 1.2 the total number of edges is therefore $\sum_{i=0}^{n-1} s(i) = f(n)$. \Box

From Observation 1.1 and Proposition 1.3 we have $f(n) \leq g(n)$. Also, by Observation 1.2 we have from [5], [6], [10] and [7] that f(n) satisfies the well-known divide-and-conquer maximin recurrence

$$f(n) = \max_{\substack{n_1+n_2=n\\n_1,n_2 \ge 1}} \left(\min(n_1, n_2) + f(n_1) + f(n_2) \right).$$
(3)

The proof that f(n) defined by (1) satisfies (3) given in [10, pages 22 - 23] is particularly short and slick.

From (3) it is evident that $g(n) \leq f(n)$ and hence we have equality, namely the following, as stated in [5].

Corollary 1.4 The maximum number of edges an induced simple graph on n vertices of a rectangular grid \mathbb{Z}^k can have is $f(n) = \sum_{i=0}^{n-1} s(i)$, the combined number of 1s in the binary expression of $0, 1, \ldots, n-1$.

REMARK: Note that the heuristic integer sequence [1, A007818] and the sequence $(f(n))_{n=0}^{\infty} = (0, 1, 2, 4, 5, 7, 9, 12, 13, 15, 17, 20, ...)$ [2, A000788] agree in the first twelve entries, but differ in the entries from and including thirteen. By Corollary 1.4 this means that the maximum number of edges of an induced graph on n vertices of an arbitrary rectangular grid \mathbb{Z}^k for $n \in \{1, 2, ..., 12\}$ can be realized in \mathbb{Z}^3 .

The divide-and-conquer maximin recurrence (3) is the best-known and most studied one, and one of the very few with an exact solution given by Observation 1.2. This is mainly since it occurs naturally when analysing worst-case scenarios in sorting algorithms [12] where both asymptotic results and some other exact solutions to more general divide-and-conquer maximin recurrences are given. The fact that f(n) defined by (1) also satisfies (3) is a consequence of a special case of the general treatment in [12]. In [11] some of the general asymptotic bounds from [12] are improved further. – The other reason the divide-and-conquer maximin recurrence (3) has been studied widely is because its solution f(n) appears as the answer to extremal combinatorial problems as in [5] where the main result is that of Corollary 1.4. In earlier articles like [8] and [9] a procedure is given on how to place the numbers $1, \ldots, 2^k$ on the vertices of the hypercube Q_k so the sum $\sum |i - j|$ over all neighbors of Q_k is minimized. Also, f satisfying (3) appears when studying the number of 1's in binary integers directly, as is done in [6]. There the main result is the presentation of tight closed lower and upper bounds for f(n), but also a description of which n_1 and n_2 adding up to n in (3) will yield the maximum of f(n). In the next section we give a new geometric characterization of the those pairs of naturals numbers adding up to n that yielding the maximum f(n) in (3).

2 Hypercubic bipartitions of an integer

In his section we give a geometric characterization of the natural numbers n_0 and n_1 summing up to n such that $f(n) = n_1 + f(n_1) + f(n_0)$. We also give a direct algebraic parametrization of such ordered pairs (n_0, n_1) . We then enumerate them for each fixed $n \in \mathbb{N}$ in the following section.

Definition 2.1 For $n \ge 2$ and $k = \lceil \lg n \rceil$, a partition $n = n_0 + n_1$ with $n_0, n_1 \ge 1$ and $n_0 \ge n_1$ is a hypercubic bipartition (HCBP) if there is an $i \in \{1, ..., k\}$ such that the hyperplane $x_i = 1/2$ splits the n points $\{\tilde{\beta}_k(0), \tilde{\beta}_k(1), ..., \tilde{\beta}_k(n-1)\}$ into two parts containing n_0 points on one side of the hyperplane and n_1 on the other.

REMARK: Of course, a partition of n is a HCBP iff there is an $i \in \{1, \ldots, k\}$ such that among the n points $\{\tilde{\beta}_k(0), \tilde{\beta}_k(1), \ldots, \tilde{\beta}_k(n-1)\}$ there are n_0 of them with *i*-th coordinate 0 and n_1 of them with *i*-th coordinate 1.¹

Our first objective is to prove the following equivalence.

Theorem 2.2 For $n \in \mathbb{N}$, a partition $n = n_0 + n_1$ with $n_0 \ge n_1$ is a HCBP if and only if $f(n) = n_1 + f(n_1) + f(n_0)$.

One direction is fairly straightforward; by Proposition 1.3 and Corollary 1.4 the *n* points $\beta_k(0), \ldots, \beta_k(n-1)$ form a maximally connected subgraph of Q_k with f(n) edges. Assume $n = n_0 + n_1$ is a HCBP. Then for some $i \in \{1, \ldots, k\}$ there are n_0 points with *i*-th coordinate 0 and n_1 points with *i*-th coordinate 1. Among the f(n) edges precisely n_1 of them are parallel to the *i*-th axis. At most $f(n_0)$ of the remaining edges connect points with *i*-th coordinate 0 and at most $f(n_1)$ connect

¹this is the reason for our change in labeling to (n_0, n_1) from (n_1, n_2) in (3).

points with *i*-th coordinate 1. By maximality of f(n) we therefore have $f(n) = n_1 + f(n_1) + f(n_0)$, and hence the following.

Observation 2.3 If $n \ge 2$ and $n = n_0 + n_1$ is a HCBP then $f(n) = n_1 + f(n_1) + f(n_0)$.

To verify the other direction, we will do so in a number of small steps. For each of them we attempt to keep our arguments as elementary as possible. The first one is obtaining an equivalent algebraic description of a HCBP of n.

Consider the $n \times k$ matrix \mathcal{B}_n where $k = \lceil \lg n \rceil$ whose *i*-th row vector is the point $\beta_k(i-1)$ for i = 1, ..., n. For each $i \in \{1, ..., k\}$ let $d_i(n)$ denote the difference between the number of 0's and the number of 1's in the *i*-th column of \mathcal{B}_n . For convenience we set $d_i(0) = 0$ for each *i*. We then have

$$d_i(n) = 2^{i-1} - |(n \mod 2^i) - 2^{i-1}| = 2^{i-1} - |n - 2^i \lfloor n/2^i \rfloor - 2^{i-1}|.$$
(4)

If the *i*-th column of \mathcal{B}_n has n_0 zeros and n_1 ones, then $n_0 - n_1 = d_i(n)$ and $n = n_0 + n_1$ is a HCBP. The converse is clear and hence we have the following characterization.

Observation 2.4 For $n \ge 2$ the partition $n = n_0 + n_1$ with $n_0, n_1 \ge 1$ and $n_0 \ge n_1$ is a HCBP if and only if $n_0 - n_1 = d_i(n)$ for some $i \in \{1, ..., k\}$ where $k = \lceil \lg n \rceil$.

Viewing *i* as fixed the graph of the map $n \mapsto d_i(n)$ has a zigzag like shape where each zig has length/period 2^i . The function d_i can easily be extended to all non-negative integers *n*.

Claim 2.5 $d_i(m) = d_i(n)$ if and only if $m \equiv \pm n \pmod{2^i}$.

Proof. The function $d_i : \mathbb{N}_0 \to \mathbb{N}_0$ is the identity on $\{0, 1, \ldots, 2^{i-1}\}$, is even, and has period 2^i . \Box

The following properties are also straightforward from the definition (4).

Claim 2.6 For $i, j \in \mathbb{N}$ and $m, n \in \mathbb{N}_0$ we have

- 1. If $i \leq j$, then $d_i(n) \leq d_j(n)$.
- 2. If $i \leq j$ and $d_j(m) = d_i(n)$, then $d_j(m) = d_i(m)$.

Claim 2.7 For $n \in \mathbb{N}_0$ and $i \in \mathbb{N}$ we have:

- 1. $d_{i+1}(2n) = 2d_i(n)$.
- 2. $d_{i+1}(2n+1) = d_i(n+1) + d_i(n)$.

For our next lemma, assume that $d_j(n+1) = d_i(n) \pm 1$ for some i, j. Since in general $d_\ell(n+1) = d_\ell(n) \pm 1$ for each ℓ , then we have a total eight cases to consider.

If $d_j(n+1) = d_i(n) + 1$ for some i and j, and $d_i(n+1) = d_i(n) - 1$ and $d_j(n+1) = d_j(n) - 1$, then $d_j(n) = d_i(n) + 2$ and $d_j(n+1) = d_i(n+1) + 2$. Therefore we have that i < j and both d_i and d_j are decreasing from n to n+1. This can only occur when $i = 1, j \ge 3$ and $n \equiv -3 \pmod{2^j}$. For these values of i and j we now have that $d_3(2n+1) = d_2(n+1) + d_2(n) = 3 = d_j(n+1) + d_i(n)$.

Similarly, if $d_j(n+1) = d_i(n) - 1$ for some i and j, and $d_i(n+1) = d_i(n) + 1$ and $d_j(n+1) = d_j(n) + 1$, then $d_i(n+1) - d_j(n+1) = d_i(n) - d_j(n) = 2$. Hence, i > j and both d_i and d_j are increasing from n to n+1. This can only occur when $j = 1, i \ge 3$ and $n \equiv 2 \pmod{2^i}$. For these values of i and j we now have that $d_3(2n+1) = 3 = d_j(n+1) + d_i(n)$.

In all the remaining cases we obtain either $d_j(n) = d_i(n)$ or $d_j(n+1) = d_i(n+1)$, and hence we obtain by Claim 2.7 the following. **Lemma 2.8** Let $i, j \in \mathbb{N}$. If $d_j(n+1) = d_i(n) \pm 1$ then $d_j(n+1) + d_i(n) = d_\ell(2n+1)$ for some $\ell \in \{3, i+1, j+1\}$.

We now have what we need to complete the more involved part of the proof of Theorem 2.2. We proceed by induction on n. The cases $n \leq 3$ being trivial, we consider even and odd cases, assume that the statement of Theorem 2.2 holds for all natural numbers less than 2n, and show it then holds for both 2n and 2n + 1.

CASE 1: If $f(2n) = 2n_1 + f(2n_1) + f(2n_0)$ where $n = n_0 + n_1$ and $n_0 \ge n_1$, then directly by (1) we get $f(n) = n_1 + f(n_1) + f(n_0)$. By induction hypothesis $n = n_0 + n_1$ is a HCBP and hence by Observation 2.4 $n_0 - n_1 = d_i(n)$ for some *i*. By Claim 2.7 we have $d_{i+1}(2n) = 2d_i(n)$, and hence $2n_0 - 2n_1 = d_{i+1}(2n)$ which shows that $2n = 2n_0 + 2n_1$ is a HCBP.

CASE 2: If $f(2n) = 2n_1 + 1 + f(2n_1 + 1) + f(2n_0 - 1)$ where $n = n_0 + n_1$ and $n_0 \ge n_1 + 1$, then by the defining recursion (1) we obtain $2f(n) = 2n_1 + f(n_1) + f(n_1 + 1) + f(n_0 - 1) + f(n_0)$. By (3) we have in general that $f(n-1) \ge n_1 + f(n_1) + f(n_0 - 1)$ and $f(n+1) \ge n_1 + 1 + f(n_1 + 1) + f(n_0)$ and hence $2f(n) + 1 \le f(n-1) + f(n+1)$ in this case. Since 2f(n) = f(2n) - n by (1), we obtain by (3) the opposite inequality and so we have here equality in both inequalities, so 2f(n) + 1 =f(n-1) + f(n+1). By Observation 1.2 this can be rewritten as s(n-1) + 1 = s(n) which means that n is odd. Also, by induction hypothesis both $n-1 = n_1 + (n_0 - 1)$ and $n+1 = (n_1 + 1) + n_0$ are HCBP and hence $(n_0 - 1) - n_1 = d_i(n-1)$ and $n_0 - (n_1 + 1) = d_j(n+1)$ for some i and j and hence $d_i(n-1) = d_j(n+1)$. By Claim 2.6 $d_\ell(n-1) = d_\ell(n+1)$ for some $\ell \in \{i, j\}$, and hence by Claim 2.5 either 2 or 2n is divisible by 2^ℓ . Since n is odd we must have $\ell = 1$ and therefore $n_0 = n_1 + 1$. As $2n_0 - 1 = 2n_1 + 1$, then $2n = (2n_1 + 1) + (2n_0 - 1)$ is a HCBP.

REMARK: We see conversely that if $n = n_0 + n_1$ is a HCBP, then by Observation 2.4 $2n_0 = n + d_i(n)$ and $2n_1 = n - d_i(n)$ for some *i*. Since $d_i(n) \equiv \pm n \pmod{2^i}$ then either n_0 or n_1 is divisible by 2^{i-1} . Therefore if a bipartition of 2n into two odd parts is a HCBP, then both parts must be equal.

CASE 3: If $f(2n+1) = 2n_1 + f(2n_1) + f(2n_0+1)$ where $n = n_0 + n_1$ and $n_0 \ge n_1$, then by (1) we get $f(n) + f(n+1) = 2n_1 + 2f(n_1) + f(n_0) + f(n_0+1)$ By (3) we have in general that $f(n) \ge n_1 + f(n_1) + f(n_0)$ and $f(n+1) \ge n_1 + f(n_1) + f(n_0+1)$, and hence in this case we have equality in both these inequalities, so $f(n) = n_1 + f(n_1) + f(n_0)$ and $f(n+1) = n_1 + f(n_1) + f(n_0+1)$, which by induction hypothesis are both HCBPs and hence $n_0 - n_1 = d_i(n)$ and $n_0 + 1 - n_1 = d_j(n+1)$ for some i and j. By Lemma 2.8 then $d_i(n) + d_j(n+1) = d_\ell(2n+1)$ for some ℓ , so $2n + 1 = 2n_1 + (2n_0 + 1)$ is a HCBP.

CASE 4: Finally, if $f(2n + 1) = 2n_1 + 1 + f(2n_1 + 1) + f(2n_0)$ where $n = n_0 + n_1$ and $n_0 \ge n_1 + 1$, then by (1) we get $f(n) + f(n + 1) = 2n_1 + 1 + f(n_1 + 1) + f(n_1) + 2f(n_0)$ By (3) we have in general that $f(n) \ge n_1 + f(n_1) + f(n_0)$ and $f(n + 1) \ge n_1 + 1 + f(n_1 + 1) + f(n_0)$, and hence in this case we have equality in both these inequalities, so $f(n) = n_1 + f(n_1) + f(n_0)$ and $f(n + 1) = n_1 + 1 + f(n_1 + 1) + f(n_0)$, which by induction hypothesis are both HCBPs and hence $n_0 - n_1 = d_i(n)$ and $n_0 - (n_1 + 1) = d_j(n + 1)$ for some i and j. By Lemma 2.8 then $d_i(n) + d_j(n + 1) = d_\ell(2n + 1)$ for some ℓ , so $2n + 1 = 2n_1 + (2n_0 + 1)$ is a HCBP. – This completes the proof of Theorem 2.2.

We conclude this section by summarizing the main results from this section.

Theorem 2.9 For $n \in \mathbb{N}$ and f the function defined in (1), TFAE:

- 1. $n = n_0 + n_1$ is a HCBP.
- 2. $f(n) = n_1 + f(n_1) + f(n_0)$ and $n_0 \ge n_1 \ge 1$.

3. $(n_0, n_1) = \left(\frac{n + d_i(n)}{2}, \frac{n - d_i(n)}{2}\right)$ for some $i \in \{1, \dots, \lceil \lg n \rceil\}$.

REMARK: In [6] a different description of nonnegative integer pairs (n_0, n_1) such that $f(n_0 + n_1) = \min(n_0, n_1) + f(n_0) + f(n_1)$ holds is given. Although there the main focus is on the entire set of such pairs, as suppose to HCBPs of each fixed $n \in \mathbb{N}$ as in Theorem 2.9, a careful reading of [6] shows that the description in Theorem 2.9 part 3 and in [6] are equivalent: plotting $\{(x, y) \in \{0, 1, \ldots\} : (x, y) = \left(\frac{n+d_i(n)}{2}, \frac{n-d_i(n)}{2}\right)$ where $n \in \mathbb{N}$ and $1 \le i \le \lceil \lg n \rceil$ yields the same set as is described in [6] minus the points on the x-axis y = 0.

3 Enumerations of HCBPs

For $n \in \mathbb{N}$ let h(n) denote the number of HCBPs $n = n_0 + n_1$ where $n_0 \ge n_1 \ge 1$. In this section we will determine the generating function $h(x) = \sum_{n\ge 0} h(n)x^n$, present a efficient recursive procedures to compute h(n), and present a formula for h(n) in terms of a special presentation of each fixed $n \in \mathbb{N}$. For this purpose it will be convenient to add the trivial partition n = n + 0 to the HCBPs of n, and so h(n) = c(n) - 1, where c(n) is the number of "non-proper" HCBPs of n in which $n_0 = n$ and $n_1 = 0$ is allowed. By definition of $d_i(n)$ from (4) we note that for each $i > \lceil \lg n \rceil$ we have $d_i(n) = n$. Hence, by Theorem 2.9

$$c(n) = \left| \left\{ \left(\frac{n + d_i(n)}{2}, \frac{n - d_i(n)}{2} \right) : i \in \mathbb{N} \right\} \right|,$$

or equivalently, c(n) is the number of distinct values of $d_i(n)$ for various $i \in \mathbb{N}$. From a geometric point of view, let $\Gamma_i = \{(n, d_i(n)) : n \in \mathbb{N}_0\}$ be the graph of the map $d_i : \mathbb{N}_0 \to \mathbb{N}_0$ for each i and let $\Gamma = \bigcup_{i \ge 1} \Gamma_i \subseteq \mathbb{N}_0^2$. If $V_n = \{(n, y) : y \in \mathbb{Z}\}$ denotes the integral vertical line in \mathbb{N}_0^2 at n, then $c(n) = |\Gamma \cap V_n|$. To obtain the generating function $c(x) = \sum_{n \ge 0} c(n)x^n$ we partition Γ into slices parallel to the x-axis as

$$\Gamma = \bigcup_{i \ge 0} (\Gamma \cap \Pi_i),$$

where $\Pi_0 = \{(x, y) \in \mathbb{N}_0^2 : 0 \le y \le 1\}$ and $\Pi_i = \{(x, y) \in \mathbb{N}_0^2 : 2^{i-1} < y \le 2^i\}$ for each $i \in \mathbb{N}$. From this we get that

$$c(n) = \left| \left(\bigcup_{i \ge 0} (\Gamma \cap \Pi_i) \right) \cap V_n \right|$$
$$= \left| \bigcup_{i \ge 0} (\Gamma \cap \Pi_i \cap V_n) \right|$$
$$= \sum_{i \ge 0} |\Gamma \cap \Pi_i \cap V_n|$$
$$= \sum_{i \ge 0} c_i(n),$$

where $c_i(n) = |\Gamma \cap \Pi_i \cap V_n|$. Since $d_i(n) \leq n$ for each fixed n and all i, we have by definition of Π_i that $\Gamma \cap \Pi_i \cap V_n = \emptyset$ for $i > \lceil \lg n \rceil$. Hence, the last sum in the above display is a finite sum. Letting $c_i(x) = \sum_{n \ge 0} c_i(n) x^n$ be the generating function corresponding to $(c_i(n))_{n \ge 0}$, we get $c(x) = \sum_{i>0} c_i(x)$. Note that $c_0(n) = 1$ for each $n \ge 0$. For $i \ge 1$ we have the following.

Lemma 3.1 Let $i \ge 1$ and $n \ge 0$. If $2^{i-1} < d_j(n) \le 2^i$, then $d_j(n) = d_{i+1}(n)$. Further $2^{i-1} < d_{i+1}(n) \le 2^i$ iff $-2^{i-1} < (n \mod 2^{i+1}) - 2^i < 2^{i-1}$.

Proof. (Sketch.) By definition, we have $0 \leq d_j(n) \leq 2^{j-1}$ for all n. So if $2^{i-1} < d_j(n) \leq 2^i$, then $i+1 \leq j$ must hold. Assuming $2^{i-1} < d_j(n) \leq 2^i$, we consider two cases. (i) If $n \mod 2^j \in \{0, 1, \ldots, 2^{j-1}\}$, then since $i \leq j-1$ we have $d_j(n) = d_{i+1}(n) = n \mod 2^{i+1}$. (ii) If $n \mod 2^j \in \{2^{j-1}+1, \ldots, 2^j-1\}$ then $d_j(n) = 2^j - (n \mod 2^j)$. Since $j-1 \leq i$ we further have by our assumption that

$$n \mod 2^j \in \{2^j - 2^i, \dots, 2^j - 2^{i-1}\} \subseteq \{2^j - 2^i, \dots, 2^j - 1\}$$

and hence $n \mod 2^{i+1} \in \{2^i, \dots, 2^{i+1} - 1\}$ as $n \mod 2^j = n \mod 2^{i+1} + 2^j - 2^{i+1}$. Therefore $d_{i+1}(n) = 2^{i+1} - n \mod 2^{i+1} = 2^j - n \mod 2^j = d_j(n)$. The rest follows from the definition of $d_{i+1}(n)$.

What the above Lemma 3.1 states is that (i) for each $n \ge 0$ we have $c_i(n) = 0, 1$, and (ii) $c_i(n) = 1$ iff $|(n \mod 2^{i+1}) - 2^i| < 2^{i-1}$. Letting $\ell = n \mod 2^{i+1}$ and writing $n = m2^{i+1} + \ell$, we consequently get for $i \ge 1$ that

$$c_{i}(x) = \sum_{\substack{0 \le n = m2^{i+1} + \ell \\ |\ell - 2^{i}| < 2^{i-1}}} x^{n}$$

$$= \sum_{m \ge 0} \left(\sum_{\ell=2^{i} - 2^{i-1} + 1}^{2^{i} + 2^{i-1} - 1} x^{m2^{i+1} + \ell} \right)$$

$$= \frac{x^{2^{i-1} + 1} (1 - x^{2^{i-1}})}{1 - x} \sum_{m \ge 0} x^{m2^{i+1}}$$

$$= \frac{x^{2^{i-1} + 1} (1 - x^{2^{i-1}})}{(1 - x)(1 - x^{2^{i+1}})}.$$

Since $c(x) = \sum_{i \ge 0} c_i(x) = c_0(x) + \sum_{i \ge 1} c_i(x)$, we obtain the following theorem.

Theorem 3.2 The generating function $c(x) = \sum_{n\geq 0} c(n)x^n$ for $(c(n))_{n\geq 0}$ is given by

$$c(x) = \frac{1}{1-x} + \sum_{i \ge 1} \frac{x^{2^{i-1}+1}(1-x^{2^{i-1}})}{(1-x)(1-x^{2^{i+1}})}.$$

Hence, since h(n) = c(n) - 1 for each $n \ge 2$, the generating function $h(x) = \sum_{n\ge 0} h(n)x^n$ for $(h(n))_{n\ge 0}$, the number of HCBPs of n, is given by

$$h(x) = \sum_{i \ge 1} \frac{x^{2^{i-1}+1}(1-x^{2^{i-1}})}{(1-x)(1-x^{2^{i+1}})}.$$

Let $n \in \mathbb{N}$ and $k = \lceil \lg n \rceil$. For $i \in \{1, \dots, k-1\}$ we have $(2^k - n) \mod 2^{i+1} = 2^{i+1} - (n \mod 2^{i+1})$ and hence $|((2^k - n) \mod 2^{i+1}) - 2^i| = |(n \mod 2^{i+1}) - 2^i|$. From this we see that if $i \in \{1, \dots, k-1\}$ then $c_i(n) = 1$ iff $c_i(2^k - n) = 1$. Additionally for i = k we have $|(n \mod 2^{k+1}) - 2^k| = 2^k - n < 2^{k-1}$ and $|((2^k - n) \mod 2^{k+1}) - 2^k| = n > 2^{k-1}$ and so $c_k(n) = 1$ and $c_k(2^k - n) = 0$. As $c_0(n) = 1$ for each n and $c_i(n) = 0$ for each i > k we have

$$c(n) = \sum_{i=0}^{k} c_i(n) = \left(\sum_{i=0}^{k-1} c_i(n)\right) + c_k(n) = \left(\sum_{i=0}^{k-1} c_i(2^k - n)\right) + 1 = \left(\sum_{i\geq 0} c_i(2^k - n)\right) + 1 = c(2^k - n) + 1,$$

which yields a recurrence for the c(n), namely

$$c(0) = c(1) = 1 \text{ and } c(n) = c(2^{\lceil \lg n \rceil} - n) + 1 \text{ for } n > 1.$$
 (5)

As the number h(n) of HCBPs of n satisfies h(n) = c(n) - 1 we have the following

Corollary 3.3 For $n \in \mathbb{N}$ the number h(n) of HCBPs of n satisfies the following determining recurrence

$$\begin{aligned} h(1) &= 0, \\ h(2) &= 1, \\ h(n) &= h(2^{\lceil \lg n \rceil} - n) + 1, \text{ for } n > 2. \end{aligned}$$

REMARK: Note that the map $n \mapsto 2^{k+1} - n$ is the reflection about the vertical line x = k in the real Euclidean plane \mathbb{R}^2 . Hence, looking at $\Gamma \subseteq \mathbb{N}_0^2 \subseteq \mathbb{R}^2$, the map $n \mapsto 2^{\lceil \lg n \rceil} - n$ is a reflection about $x = \lceil \lg n \rceil - 1$, 2 to the power of which is the largest power of 2 strictly less than n. From the shape of Γ it is clear that we have a bijection

$$(\Gamma \cup V_n) \setminus \{(n,n)\} \to \Gamma \cup V_{2^{\lceil \lg n \rceil} - n}$$

given by $(n, y) \mapsto (2^{\lceil \lg n \rceil} - n, y)$. Hence, the recursion in (5) (and therefore Corollary 3.3) is also evident from this geometrical perspective.

If $n \in \mathbb{N}$ and $k = \lceil \lg n \rceil \ge 2$, then $2^{k-1} < n \le 2^k$ and hence $0 \le 2^k - n < 2^{k-1}$. Hence both c(n) and h(n) can be computed in at most $\lceil \lg n \rceil - 1$ steps², provided that $n \ge 3$. In fact, if $n = a_k = (2^{k+1} + (-1)^k)/3$, then by (5) $c(n) = c(a_k) = c(a_{k-1}) + 1$, and so $c(n) = k = \lceil \lg n \rceil$, and is obtained in exactly $\lceil \lg n \rceil - 1$ steps with the recursion in (5).

Observation 3.4 For $n \in \mathbb{N}$ we have

- 1. $n \in \{1, 2\}$ implies $c(n) = \lceil \lg n \rceil + 1$ and hence $h(n) = \lceil \lg n \rceil$.
- 2. $n \ge 3$ implies $c(n) \le \lceil \lg n \rceil$ and hence $h(n) \le \lceil \lg n \rceil 1$, and equality holds for infinitely many $n \ge 3$.

Ideally we would like to develop an explicit formula for c(n) and hence h(n) in terms of n. We will conclude this section by the next best thing; a formula in terms of a special representation of n, similar to the one in Observation 1.2 for f(n).

²i.e. arithmetic operations

Consider $n \in \mathbb{N}$ in its binary representation. Consider each maximal string of 1s in this representation, except the last one if it is single (that is, if n is of the form $n = 2^{\alpha}(4m + 1)$.) Rewrite each of these strings as a difference of two powers of twos; $2^{\alpha} + 2^{\alpha-1} + \cdots + 2^{\beta} = 2^{\alpha+1} - 2^{\beta}$. In this way we obtain from the binary representation of n a representation of n as a finite alternating sum of powers of two's

$$n = 2^{\alpha_1} - 2^{\alpha_2} + \dots + (-1)^{\ell-1} 2^{\alpha_\ell}, \tag{6}$$

where

$$\alpha_1 > \alpha_2 > \dots > \alpha_{\ell-1} > \alpha_\ell + 1. \tag{7}$$

Definition 3.5 A representation of $n \in \mathbb{N}$ as (6) where the exponents satisfy (7) is called an alternating binary representation (ABR) of n.

An ABR of $n \in \mathbb{N}$ has the following property.

Lemma 3.6 If $(\alpha_1, \ldots, \alpha_\ell)$ determines an ABR of $n \in \mathbb{N}$, then $\alpha_1 = \lceil \lg n \rceil$ and $\alpha_i = \lceil \lg((-1)^{i-1}(n-2^{\alpha_1}+2^{\alpha_2}+\cdots+(-1)^{i-1}2^{\alpha_{i-1}})) \rceil$ for each $i \in \{2, \ldots, \ell\}$. In particular, the ABR of n is unique and $(\alpha_2, \ldots, \alpha_\ell)$ determines the ABR of $2^{\alpha_1} - n$.

Proof. Let $n = 2^{\alpha_1} - 2^{\alpha_2} + \dots + (-1)^{\ell-1} 2^{\alpha_\ell}$ where $\alpha_1 > \alpha_2 > \dots > \alpha_{\ell-1} > \alpha_\ell + 1$. We proceed by induction on ℓ : the statement is clear for $\ell = 1$ as n is then a power of 2.

Let $\ell \geq 2$. In this case we have $2^{\alpha_1} > n$. If $\ell = 2$, then $n = 2^{\alpha_1} - 2^{\alpha_2}$ where $\alpha_1 > \alpha_2 + 1$ and hence $n > 2^{\alpha_2+1} - 2^{\alpha_2} = 2^{\alpha_2}$. So we have here that $2^{\alpha_2} < n < 2^{\alpha_1}$ and therefore $\alpha_1 = \lceil \lg n \rceil$. Since $2^{\alpha_1} - n = 2^{\alpha_2}$ the lemma holds in this case.

If $\ell \geq 3$, then $n = 2^{\alpha_1} - 2^{\alpha_2} + n''$ where n'' > 0 and hence $n > 2^{\alpha_1} - 2^{\alpha_2} \geq 2^{\alpha_2}$. So we have here that $2^{\alpha_2} < n < 2^{\alpha_1}$ and therefore $\alpha_1 = \lceil \lg n \rceil$. Since $(\alpha_2, \ldots, \alpha_\ell)$ determines an ABR of $n' = 2^{\alpha_1} - n$, the lemma follows by induction on ℓ .

We now have a formula for c(n) and h(n) in terms of the ABR of n.

Theorem 3.7 For $n \in \mathbb{N}$ let $n = 2^{\alpha_1} - 2^{\alpha_2} + \cdots + (-1)^{\ell-1} 2^{\alpha_\ell}$ be the ABR of n. Then

$$c(n) = \begin{cases} \ell & \text{if } n \text{ is odd } (\alpha_{\ell} = 0), \\ \ell + 1 & \text{if } n \text{ is even } (\alpha_{\ell} \ge 1). \end{cases}$$

Consequently, for the number of HCBP h(n) we have

$$h(n) = \begin{cases} \ell - 1 & \text{if } n \text{ is odd } (\alpha_{\ell} = 0), \\ \ell & \text{if } n \text{ is even } (\alpha_{\ell} \ge 1). \end{cases}$$

Proof. By (5) and Corollary 3.3 the statement is clearly true for any power of two $n = 2^{\alpha}$. The rest follows by Lemma 3.6 and induction on n, and (5) and Corollary 3.3

We can extract alternative recursions for c(n) and h(n) from the above Theorem 3.7, different from the ones given in (5) and Corollary 3.3.

Note that for an even n, the ABR of n/2 is obtained by subtracting 1 from each of the α_i in the ABR of n. If n = 4m, then n/2 = 2m is still even and we have by Theorem 3.7 that

c(4m) = c(2m). If n = 4m + 2, then n/2 = 2m + 1 is odd and we obtain by Theorem 3.7 that c(4m + 2) = c(2m + 1) + 1.

For odd n, we must consider the following cases. If n = 8m + 1 (resp. n = 8m + 7), then the ABR of 4m + 1 (resp. 4m + 3) is obtained by subtracting 1 from all α_i except the last one α_ℓ in the ABR of 8m + 1 (resp. 8m + 7). As all are odd numbers, we we have by Theorem 3.7 that c(8m + 1) = c(4m + 1) (resp. c(8m + 7) = c(4m + 3).) If n = 8m + 3, then the ABR of 4m + 1 is obtained by subtracting 1 from all α_i where $i \in \{1, \ldots, \ell - 2\}$ in the ABR of 8m + 3, and replacing the last two summands $2^2 - 2^0$ for $i \in \{\ell - 1, \ell\}$ by 2^0 . As both 8m + 3 and 4m + 1 are odd, and the latter has on fewer terms in its ABR, we have c(8m + 3) = c(4m + 1) + 1. Finally, if n = 8m + 5, then the ABR of 4m + 3 is obtained by subtracting 1 from all α_i where $i \in \{1, \ldots, \ell - 2\}$ in the ABR of 8m + 5, then removing -2^2 for $i = \ell - 1$ and replacing 2^0 by -2^0 for $i = \ell$. Again, as both 8m + 5 and 4m + 3 are odd, and the latter has on fewer terms in its ABR, we have terms in its ABR, we have c(8m + 5) = c(4m + 3) = c(4m + 3

Corollary 3.8 For $n \in \mathbb{N}$ the number h(n) of HCBPs of n is determined by

$$(h(n))_{n=1}^{8} = (0, 1, 1, 1, 2, 2, 1, 1)$$

and the following recurrence

$$h(n) = \begin{cases} h(n/2) & \text{if } n \equiv 0 \pmod{4}, \\ h(n/2) + 1 & \text{if } n \equiv 2 \pmod{4}, \\ h((n+1)/2) & \text{if } n \equiv 1 \pmod{8}, \\ h((n-1)/2) + 1 & \text{if } n \equiv 3 \pmod{8}, \\ h((n+1)/2) + 1 & \text{if } n \equiv 5 \pmod{8}, \\ h((n-1)/2) & \text{if } n \equiv 7 \pmod{8}. \end{cases}$$

REMARK: There is a strong resemblance between c(n) and h(n) on one hand, and s(n), the number of digits in the binary representation of n from (2), on the other; firstly s(n) satisfies the recursion $s(n) = s(n - 2^{\lfloor \lg n \rfloor}) + 1$ with s(0) = 0, a very similar recursion to the one in (5). These recursions are both obtained "from the top", or "from the left", in the sense that we consider what happens when we remove the first power of 2 in the usual binary representation of n and in the ABR of nrespectively. On the other hand the recursion in (2) and Corollary 3.8 are both obtained "from behind", or "from the right", by considering the removal of the last power of 2 in the usual binary representation of n and in the ABR of n respectively.

Note that by Observation 3.4 we have that $2 \le c(n) \le \lceil \lg n \rceil$ and $1 \le h(n) \le \lceil \lg n \rceil - 1$ for every $n \ge 3$. We conclude this section by a sharpening of this about the number of HCBPs of n with $\lceil \lg n \rceil = k$ given.

Proposition 3.9 For $k \ge 2$ and $\ell \in \{1, \ldots, k-1\}$ let $\mathcal{H}_k(\ell) = \{n : \lceil \lg n \rceil = k \text{ and } h(n) = \ell\}$. Then

$$|\mathcal{H}_k(\ell)| = 2\binom{k-2}{\ell-1}.$$

In particular for $k = \lceil \lg n \rceil$ we have in the extreme cases that

1. h(n) = 1 iff $n \in \{2^k - 1, 2^k\}$, 2. h(n) = k - 1 iff $n \in \{(2^{k+1} + (-1)^k)/3, (2^{k+1} + (-1)^k)/3 + (-1)^k\}$. *Proof.* Assume that $\lceil \lg n \rceil = k$ and hence $2^{k-1} < n \le 2^k$. If $(\alpha_1, \ldots, \alpha_i)$ determine the ABR of n, then $\alpha_1 = k$ is determined.

FIRST CASE: If n is even, then $\alpha_i \geq 1$ and hence the distinct decreasing exponents $\alpha_2, \dots, \alpha_{i-1}, \alpha_i + 1$ are determined by an (i-1)-subset of $\{2, \dots, k-1\}$ of which there are exactly $\binom{k-2}{i-1}$. By Theorem 3.7 the number of even $n \in \mathcal{H}_k(\ell)$ is therefore given by $\binom{k-2}{\ell-1}$.

SECOND CASE: If n is odd, then $\alpha_i = 0$ is determined and hence the distinct decreasing exponents $\alpha_2, \dots, \alpha_{i-1}$ are determined by an (i-2)-subset of $\{2, \dots, k-1\}$ of which there are exactly $\binom{k-2}{i-2}$. By Theorem 3.7 the number of odd $n \in \mathcal{H}_k(\ell)$ is therefore given by $\binom{k-2}{(\ell+1)-2} = \binom{k-2}{\ell-1}$. Hence $|\mathcal{H}_k(\ell)| = 2\binom{k-2}{\ell-1}$, which completes the first part of the Proposition.

By Corollary 3.3 we clearly have $h(2^k) = h(2^k - 1) = 1$, and by the first part there are precisely two numbers n with $\lceil \lg n \rceil = k$ with h(n) = 1.

Finally, if $n = a_k = (2^{k+1} + (-1)^k)/3 = 2^k - 2^{k-1} + \dots + (-1)^{k-2}2^2 + (-1)^{k-1}2^0$, then we have seen right above Observation 3.4 that h(n) = c(n) - 1 = k - 1. Also, for an even $n = b_k = a_k + (-1)^k = 2^k - 2^{k-1} + \dots + (-1)^{k-3}2^3 + (-1)^{k-2}2^2$, then we have by Corollary 3.3 that $h(b_k) = h(b_{k-1}) + 1 = h(b_2) + k - 2 = h(4) + k - 2 = k - 1$. Since by the first part there are at most two numbers n with $\lceil \lg n \rceil = k$ and h(n) = k - 1, this completes the proof.

Acknowledgments

Sincere thanks to the anonymous referees for ...

References

- [1] Sequence A007818, The On-Line Encyclopedia of Integer Sequences (OEIS): http://oeis.org/A007818.
- [2] Sequence A000788, The On-Line Encyclopedia of Integer Sequences (OEIS): http://oeis.org/A000788.
- [3] Michael Trott: The Mathematica GuideBook for Programming, *Springer Verlag*, New York, (2004). http://www.mathematicaguidebooks.org/.
- [4] David Tall: The blancmange function. Continuous everywhere but differentiable nowhere, Math. Gaz., 66:no. 435, 11–22, (1982).
- [5] Sergiu Hart: A note on the edges of the n-cube, Discrete Math., 14:no. 2, 157 163, (1976).
- [6] M. D. McIlroy: The number of 1's in binary integers: bounds and extremal properties, SIAM J. Comput., 3, 255 – 261, (1974).
- [7] Hubert Delange: Sur la fonction sommatoire de la fonction "somme des chiffres". (French), Enseignement Math., (2) 21, no. 1, 31 – 47, (1975).
- [8] L. H. Harper: Optimal assignments of numbers to vertices, J. Soc. Indust. Appl. Math., 12, 131 – 135, (1964).
- [9] A. J. Bernstein: Maximally connected arrays on the n-cube, SIAM J. Appl. Math., 15, 1485 - 1489, (1967).

- [10] Daniel H. Greene; Donald E. Knuth: Mathematics for the analysis of algorithms. Third edition, Progress in Computer Science and Applied Logic, 1. *Birkhuser Boston, Inc.*, Boston, MA, (1990).
- Biing-Feng Wang: Tighter bounds on the solution of a divide-and-conquer maximin recurrence, J. Algorithms, 23, no. 2, 329 – 344, (1997).
- [12] Zhiyuan Li; Edward M. Reingold: Solution of a divide-and-conquer maximin recurrence, SIAM J. Comput., 18 no. 6, 1188 – 1200, (1989).
- [13] Richard Bellman; Harold N. Shapiro: On a problem in additive number theory, Ann. of Math., (2) 49, 333 – 340, (1948).

June 27, 2011