On a Hamilton walk along the vertices of the hypercube

Louis ten Bosch, Matthijs Coster

May 5, 2003

1 Abstract

In this paper, we study the existence of a special Hamilton path along the vertices of the (real) hypercube. This path is described by an iteration of a mapping, which is the composition of a real linear transformation followed by a quantization step that preserves the sign of each coordinate. We will proof that such paths exist for each dimension and show a number of interesting and intriguing properties of these paths, related to Fibonacci numbers and Fibonacci matrices.

Keywords: Hypercube, Hamilton path, Fibonacci matrices.

2 Introduction

In this paper, we discuss the relation between certain N-dimensional (real) integer matrices with Fibonaccian properties on the one hand, and Hamilton paths on N-dimensional hypercubes on the other. A possible connection between these concepts goes via computational models known as Artificial Neural Networks (ANNs). In the last three decades, ANNs have received considerable interest as part of the much wider research field of Artificial Intelligence. In this research field, one aims at the modelling of process of complex input processing such as (human) learning, remembering, reasoning, and language comprehension. Originally, ANNs can be considered as a representation of the human brain with the characteristics of being relatively tolerant against faulty input, parallelism, and 'adaptive' behavior with respect to changes in the input. While the actual function of the human brain is still far from understood, the mathematical/numerical theory underlying ANNs has evolved considerably since the 1980s. The construction of a reasonable or valid ANN has its roots in approximation theory, computer science, optimization research and physics (statistical mechanics) (multi-layer perceptrons, Minsky & Papert (1969), backward propagation (Rumelhart, 1975; Rumelhart & McClelland, 1986), energy minimization (Hopfield, 1982-4)). Nowadays, four important fields of applications can be distinguished: classification, association, optimization, and self-organization. ANNs as classifiers are widely used in areas such as automatic speech recognition, handwriting recognition, and image processing.

The general approach to ANNs is purely mathematical. The basic building block of an ANN is a unit (sometimes called neuron or node). An ANN can be regarded as a connected graph G with units as vertices. This graph G is weighted and directed but not necessarily complete. Each vertex attains a (real) value. Dynamic behavior is introduced by having the value of a vertex updated on the basis of neighboring vertices, for example the set of vertices w such that $(w, v) \in E(G)$, the set of edges of G. (This set may include the vertex v itself.) A well-known example of such an 'activation' function reads

$$v_n^{\text{updated}} = \operatorname{sgn}\left(\sum_i w_{i\to n} v_i\right).$$

in which v_n denotes the value of node n, $\operatorname{sgn}(x)$ denotes the sign of x (i.e. $\operatorname{sgn}(x) = 0$ if $x = 0$, else $\operatorname{sgn}(x) = x/|x|$, and the sum is taken over all nodes v_i such that $(v_i, v_n) \in E(G)$. The weights $w_{i\to n}$ denote the weight of the edge (v_i, v_n). This particular activation function is sometimes referred to as 'Heaviside function'.

Any activation function can be considered as a 'locally' defined receipt for the update of the value of a node v depending on the values of the nodes in a set (which depends on v). After updating one particular node, it depends on the update strategy which node is to be updated next. Several updating schemes have been proposed in the literature, varying from unchangeable schemes or random visiting schemes to various types of adaptive schemes. In this paper, we will only deal with the case of a simultaneous update of all node values in the network. Suppose a network consists of d nodes. Given the activation function presented above, the set of values of all the nodes defines a d-dimensional vector with components in $\{-1, 0, 1\}$. A simultaneous update of all the nodes naturally corresponds to one mapping F from the d-

dimensional set $\{-1,0,1\}^d$ to itself. F can be considered as a composition of two mappings, viz. a linear mapping, followed by the vector-based extension of $\mathrm{sgn}(x)$. In summary,

$$F : x \to \mathrm{sgn} \bullet Ax = \mathrm{sgn}(Ax)$$

in which A denotes a real d-by-d matrix, and x an element in $\{-1,0,1\}^d$.

Evidently, for fixed A, iteration of F starting from a vertex v defines a closed path, the limiting cycle, of which the length only depends on A and v. Moreover, for each limiting cycle C, $-C$ is also a limiting cycle (possibly coinciding with C).

We now look for a matrix A with some specific properties. Let H^d denote the (real) d-dimensional hypercube with points with coordinates in $\{-1,1\}$. We constrain A such that that $F : x \to \mathrm{sgn}(Ax)$ describes a mapping from H^d to itself.

Main proposition. There exists an integer matrix A_d such that iteration of $F : x \to \mathrm{sgn}(A_d x)$ describes a Hamilton path in H^d.

Remark. This result was found by first considering the low-dimensional real case. The first step was to construct a possible (feasible) Hamilton path in H^d for $k = 3,4$. In the second step, a real d-by-d matrix was found in a numerical way, The third step consisted in constraining this real matrix to an integer solution.

In order to describe the Hamilton paths in dimension d we define the $d \times 2^d$ matrices P_d. P_1 is defined by

$$P_1 = (-1 \quad 1).$$

By induction we define

$$P_{i+1} = \begin{pmatrix} P_i & -P_i \\ -1 \cdots -1 & 1 \dots 1 \end{pmatrix}$$

The associated matrices A_d are defined by

$$A_1 = (-1)$$

$$A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

A_{d+1} can be constructed from A_{d-1} and A_d by

$$A_{d+1} = A_d + A_{d-1} + J_{d+1}.$$

Here we pad the $d-1\times d-1$–dimensional and $d\times d$–dimensional matrices A_{d-1} and A_d with zero's to $d+1\times d+1$–dimensional matrices. Further we define J by

$$J_n = \begin{pmatrix} O & 0 & -j \\ 0 & 1 & -1 \\ k^{\mathrm{T}} & 1 & n-2 \end{pmatrix}.$$

where j is the $n-2$–dimensional column 1–vector, and k is the $n-2$–dimensional alternating ± 1 row vector with final component -1.

Example ($d=9$).

$$A_9 = \begin{pmatrix}
-13 & -21 & -13 & -8 & -5 & -3 & -2 & -1 & -1 \\
21 & 13 & -13 & -8 & -5 & -3 & -2 & -1 & -1 \\
-13 & 13 & 21 & -8 & -5 & -3 & -2 & -1 & -1 \\
8 & -8 & 8 & 21 & -5 & -3 & -2 & -1 & -1 \\
-5 & 5 & -5 & 5 & 18 & -3 & -2 & -1 & -1 \\
3 & -3 & 3 & -3 & 3 & 14 & -2 & -1 & -1 \\
-2 & 2 & -2 & 2 & -2 & 2 & 11 & -1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & 7 & -1 \\
-1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 7
\end{pmatrix}$$

In general, the right upper part of A_d contains (from right to left) negated versions of the Fibonacci sequence, while the left bottom part is a alternating mirrored version of the right-upper part. All column sums are -1 or $+1$; moreover, they are alternating, with the sum of the rightmost column entries equalling -1.

The proposition as stated above is equivalent to the proposition that for each integer $d > 0$ $F_d = \mathrm{sgn} \bullet A_d$ acts as a (cyclic) shift on the set P_d (interpreted as a sequence of column vectors). This will be proven in the next section.

3 Proof

The proof is by induction to the dimension. For $d = 1$ and 2, the proof is trivial, so we assume $d > 2$. We will check that F_{d+1} acts as a shift on the set P_{d+1}. Let p_{di} denote the i–th column in P_d, which corresponds to the i–th point of the d–dimensional hypercube which will be visited during the walk. For simplicity for $0 \le i \le 2^d$, we denote $p_{d,2^d+i} = p_{d,i}$ and for $0 \le i \le 2^{d-1}$

we denote $p_{d-1,k2^{d-1}+i} = p_{d-1,i}$. So we have to proof

$$F_{d+1}p_{d+1,i} = p_{d+1,i+1}, \text{ for } 0 \leq i < 2^{d+1}.$$

During the proof we will denote by $c_{d,i} = A_d p_{d,i}$ and $j_{d,i} = J_d p_{d,i}$. Notice that the coordinates of $f_{d,i}$ are integral and non-zero and $p_{d,i+1} = \operatorname{sgn}(c_{d,i})$ and the first $d-1$ coordinates of $j_{d,i}$ are in absolute value exactly equal to 1. Further notice that

$$p_{d+1,i} = \begin{cases} \begin{pmatrix} p_{d-1,i} \\ -1 \\ -1 \end{pmatrix} & \text{for} \quad 0 \leq i < 2^{d-1} \\ \begin{pmatrix} p_{d-1,i-2^{d-1}} \\ 1 \\ -1 \end{pmatrix} & \text{for} \quad 2^{d-1} \leq i < 2^{d} \\ \begin{pmatrix} p_{d-1,i-2^{d}} \\ 1 \\ 1 \end{pmatrix} & \text{for} \quad 2^{d} \leq i < 3 \cdot 2^{d-1} \\ \begin{pmatrix} p_{d-1,i-3\cdot 2^{d-1}} \\ -1 \\ 1 \end{pmatrix} & \text{for} \quad 3 \cdot 2^{d-1} \leq i < 2^{d+1} \end{cases}$$

Now we use the definition $A_{d+1} = A_d + A_{d-1} + J_{d+1}$ (after appropriate zero padding). We obtain:

$$A_{d+1}p_{d+1,i} = A_d p_{d+1,i} + A_{d-1}p_{d+1,i} + J_{d+1}p_{d+1,i} = c_{d,i} + c_{d-1,i} + j_{d+1,i}.$$

For the first $d-1$ coordinates the proof is easy, since $\operatorname{sgn}(c_{d,i}) = \operatorname{sgn}(c_{d-1,i})$ and $|j_{d+1,i}| = 1$ cannot influence the sign. Therefore $p_{d+1,i+1} = \operatorname{sgn}(c_{d,i})$. The correctness for the last ($d+1^{\text{th}}$) coordinate is straightforward since the last row of A_{d+1} (equals $((-1)^d, \ldots, -1, 1, d-2)$), only changes sign if $p_{d+1,i} = \pm((-1)^d, \ldots, -1, 1, -1)$ which is the case if $i = 2^d$ or $i = 2^{d+1}$. $((-1)^d, \ldots, -1, 1, d-2)$. The correctness for the one but last (d^{th}) coordinate is comparison. We find $p_{d+1,i} = \pm((-1)^d, \ldots, -1, 1, 1)$, which implies that $i = 2^{d-1}$ or $i = 3 \cdot 2^{d-1}$.

Proposition 2 *Write A_d by A. Let $v_1^{\mathrm{T}} = (1, 1, 0, \ldots, 0)$, $v_2^{\mathrm{T}} = (1, -1, -2, 0, \ldots, 0)$, $v_3^{\mathrm{T}} = (1, -1, 1, 3, 0, \ldots, 0)$, $\ldots$, $v_{n-1}^{\mathrm{T}} = (1, -1, \ldots, (-1)^n, (-1)^n \cdot (n-1))$, $v_n^{\mathrm{T}} = (1, -1, \ldots, (-1)^n, (-1)^{n+1})$, $A_n^{\mathrm{T}} A_n$ has n integral square eigenvalues. The related eigenvectors are $v_1, \ldots, v_n$.*

The square roots of the eigenvalues are given in the following table:

dim A	1	2	3	4	5	6	7	8	9
1	1								
2	1	1							
3	2	2	1						
4	3	3	3	1					
5	5	5	4	4	1				
6	8	8	7	5	5	1			
7	13	13	11	9	6	6	1		
8	21	21	18	14	11	7	7	1	
9	34	34	29	23	17	13	8	8	1

Consequence Consider the singular value decomposition of A:

$$A = USV^{\mathrm{T}}$$

with U and V orthonormal and S a diagonal matrix. Then

$$AA^{\mathrm{T}} = USV^{\mathrm{T}}VS^{\mathrm{T}}U^{\mathrm{T}} = US^2U^{\mathrm{T}}$$

so U is the matrix of eigenvectors of AA^{T} corresponding to the squared eigenvalues of A. (Likewise V is the matrix of eigenvectors of $A^{\mathrm{T}}A$.)

4 Proof

Let $w_1^{\mathrm{T}} = (1, -1, 0, \ldots, 0)$, $w_2^{\mathrm{T}} = (1, 1, -2, 0, \ldots, 0)$, $w_3^{\mathrm{T}} = (1, 1, 1, -3, 0, \ldots, 0)$, $\ldots$, $w_{n-1}^{\mathrm{T}} = (1, \ldots, 1, -n+1)$, $w_n^{\mathrm{T}} = (1, \ldots, 1)$. We will proof that $A_n v_i = (-1)^i \cdot \lambda_{ni} w_i$ and $A_n^{\mathrm{T}} w_i = (-1)^i \cdot \lambda_{ni} v_i$. The proof of the latter statement follows immediately from the first statement by considering the transposition. The numbers λ_{ni} are given by the table. For $i < n-1$ we have $\lambda_{ni} = \lambda_{n-1,i} + \lambda_{n-2,i}$, $\lambda_{n,n-1} = n-1$ and $\lambda_{nn} = 1$. Our proof will be done by induction. For $n = 1, 2$ the proof is trivial. We may assume that the statement hold for all values up to n. For dimension $n+1$, we distinguish three cases: $1 \leq i \leq n-1$, $i = n$ and $i = n+1$.

First case This case is obvious. By induction we know that $A_{n-1} v_i = (-1)^i \cdot \lambda_{n-1i} w_i$, $A_n v_i = (-1)^i \cdot \lambda_{ni} w_i$ and furthermore $J_n v_i = 0$. Using the splitting of A_{n+1} we find $A_{n+1} v_i = (-1)^i \cdot (\lambda_{n-1i} + \lambda_{ni}) w_i$.

Second case If $i = n$, $A_{n-1} v_n = (-1)^i \cdot \lambda_{n-1i} w_i$, $A_n v_n = (-1)^i \cdot \lambda_{ni} w_i$ and furthermore $J_n v_i = 0$. !!!This has to be filled in later.

Third case By induction we know that $A_{n-1}v_{n+1} = A_{n-1}v_{n-1} = (-1)^{n-1}(1,1,\ldots,1,0,0)$, $A_n v_{n+1} = (-1)^{n-1}(1,1,\ldots,1,0)$ and furthermore $J_n v_{n+1} = (-1,-1,\ldots,-1,0,x)$.
Using the splitting of A_{n+1} we find $A_{n+1}v_i = (-1)^i \cdot (\lambda_{n-1i} + \lambda_{ni})w_i$.

5 Future research

5.1 Uniqueness of P

We wonder whether two or more columns of P can be interchanged while keeping a cycle of length 2^d. We found that for $d = 3$ other Hamilton pathes:
If

$$P = \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

For this Hamilton path we find the matrix

$$A = \begin{pmatrix} -1 & 1 & -1 \\ -1 & -1 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

For higher dimensions no other Hamilton pathes are known.

5.2 Including 0

What happens if we use $H = \{-1, 0, 1\}$ instead of $H = \{-1, 1\}$. We found for dimension 2:

$$P = \begin{pmatrix} 1 & 1 & 1 & 0 & -1 & -1 & -1 & 0 \\ 1 & 0 & -1 & -1 & -1 & 0 & 1 & 1 \end{pmatrix}$$

and the matrix

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

for dimension 3:

$$P = \begin{pmatrix} 1 & 0 & 1 & -1 & 1 & -1 & 0 & -1 & -1 & -1 & 0 & -1 & 1 & -1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & -1 & 0 & -1 & -1 & -1 & -1 & -1 & 0 & -1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 \end{pmatrix}$$

For this Hamilton path we find the matrix

$$A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & 1 & -1 \\ 1 & 1 & 2 \end{pmatrix}$$

In dimension 4 a Hamilton path of length 38 was found for the matrix:

$$A = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 3 & 1 & 2 & -1 \end{pmatrix}$$