Discrete Mathematics 50 (1984) 15–28
North-Holland

INCREASING SEQUENCES WITH NONZERO BLOCK SUMS AND INCREASING PATHS IN EDGE-ORDERED GRAPHS

A.R. CALDERBANK and F.R.K. CHUNG

Mathematics and Statistics Research Center, Bell Laboratories, Murray Hill, NJ 07974, USA

D.G. STURTEVANT

Department of Applied Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 13 September 1982

Consider the maximum length $f(k)$ of a (lexicographically) increasing sequence of vectors in $GF(2)^k$ with the property that the sum of the vectors in any consecutive subsequence is nonzero modulo 2. We prove that $\frac{23}{48} \cdot 2^k \leq f(k) \leq (\frac{1}{2}+o(1))2^k$.

A related problem is the following. Suppose the edges of the complete graph K_n are labelled by the numbers $1, 2, \ldots, \binom{n}{2}$. What is the minimum $\alpha(n)$, over all edge labellings, of the maximum length of a simple path with increasing edge labels? We prove that $\alpha(n) \leq (\frac{1}{2}+o(1))n$.

1. Introduction

A sequence of vectors $a_1, \ldots, a_t$ in $\mathrm{GF}(2)^k$ is said to be an *f-sequence* if it is increasing lexicographically and if the (modulo 2) sum of the vectors in any consecutive subsequence is nonzero (that is, $\sum_{i=r}^{s} a_i \neq 0$ for any $1 \leq r < s \leq t$). Let $f(k)$ denote the maximum length of an *f-sequence* in $\mathrm{GF}(2)^k$. Similarly, an increasing sequence of vectors $b_1, \ldots, b_t$ in $\mathrm{GF}(2)^k$ is said to be a *g-sequence* if the sum of the vectors in any consecutive subsequence of even length is nonzero (that is, $\sum_{i=r}^{r+2s-1} b_i \neq 0$ for any $1 \leq r < r+2s-1 \leq t$). Let $g(k)$ denote the maximum length of a *g-sequence* in $\mathrm{GF}(2)^k$. A little calculation reveals that $f(2)=2$, $f(3)=5$, $f(4)=10$, $g(2)=3$, $g(3)=6$ and $g(4)=12$. We shall prove that

$$\tfrac{23}{48} \cdot 2^k \leq f(k) \leq g(k) \leq (1+o(1))2^{k-1}.$$

A related problem concerning edge ordered graphs was first raised by Chvátal and Komlós in [2]: Suppose the edges of the complete graph K_n are labelled by the numbers $1, 2, \ldots, \binom{n}{2}$. What is the minimum $\alpha(n)$, over all edge labellings, of the maximum length of a simple path with increasing edge labels? Graham and Kleitman [3] proved that

$$\tfrac{1}{2}(\sqrt{4n-3}-1) < \alpha(n) < \tfrac{3}{4}n.$$

The upper bound has been improved to $\frac{7}{12}n$ by Alspach, Heinrich, and Graham

[1]. We shall prove that

$$\alpha(n) < (\tfrac{1}{2} + o(1))n$$

using the results on f- and g-sequences.

2. Increasing sequences

Given $\varepsilon = 0$ or 1 and $\nu = (\nu_{k-1}, \ldots, \nu_0) \in GF(2)^k$ define $\varepsilon\nu \in GF(2)^{k+1}$ by $\varepsilon\nu = (\varepsilon, \nu_{k-1}, \ldots, \nu_0)$. Given a sequence $A = (a_i)_{i=1}^t$ of vectors in $GF(2)^k$ let εA ($\varepsilon = 0$ or 1) denote the sequence $(b_i)_{i=1}^t$ where $b_i = \varepsilon a_i$. If A is an f-sequence of length $f(k)$, then setting $A = (0B, 1C)$ we observe that B is an f-sequence, C is a g-sequence and

$$f(k) \leq f(k-1) + g(k-1). \tag{2.1}$$

If A is a g-sequence of length $g(k)$, then setting $A = (0B, 1C)$ we observe that B and C are g-sequences and

$$g(k) \leq 2g(k-1). \tag{2.2}$$

(Notice that if A is an f- or g-sequence, then it is not always true that $(0A, 1A)$ is an f- or g-sequence.)

Therefore define an *f^*-sequence* in $GF(2)^k$ to be an increasing sequence of vectors $(a_i)_{i=1}^t$ such that $\sum_{i=r}^{r+s} a_i \neq 0$ for any $1 \leq r \leq t$, $0 \leq s < t$ (the subscripts are to be taken modulo t). Let $f^*(k)$ denote the maximum length of an f^*-sequence in $GF(2)^k$. Examples of f^*-sequences in $GF(2)^4$ and $GF(2)^5$ are given in (2.3) and (2.4) below.

0001		00001	01111	
0010		00100	10000	
0100		00110	10001	
0101	(2.3)	00111	10010	(2.4)
0110		01001	10100	
1000		01010	11001	
1001		01100	11101	
1111		01101	11111	

Clearly

$$f^*(k) \leq f(k) \leq g(k) \quad \text{for all } k. \tag{2.5}$$

If $(a_i)_{i=1}^t$ is an f^*-sequence, then the $2t-1$ sums of the form $\sum_{i=1}^r a_i$ $(1 \leq r \leq t-1)$ and $\sum_{i=s}^t a_i$ $(1 \leq s \leq t)$ are distinct and non-zero. (If 2 sums were equal, then $\sum_{i=j}^{j+k} a_i = 0$ for some $1 \leq j \leq t$, $0 \leq k < t$, contrary to the definition of an f^*-sequence.) Hence

$$f^*(k) \leq 2^{k-1}, \tag{2.6}$$

and the f^*-sequences given in (2.3) and (2.4) have maximal length.

Theorem 2.7. $2^{-k}f^*(k) \geqslant \frac{23}{48}$, *for all* k.

Proof. Let $A = (a_1, \ldots, a_t)$ be an f^*-sequence with the property that $a_1 + a_2 < a_3$. Then $A' = (a_1 + a_2, a_3, \ldots, a_t)$ is also an f^*-sequence. If t is odd, then it is straightforward to check that $(0A, 1A)$ is an f^*-sequence and $0a_1 + 0a_2 < 0a_3$. If t is even, then $(0A, 1A')$ is an f^*-sequence and again $0a_1 + 0a_2 < 0a_3$. (When t is even the only fact preventing $(0A, 1A)$ from being an f^*-sequence is that the sum of all the terms is zero.) Observe that the first 3 terms of the f^*-sequence (2.4) are 00001, 00100, and 00110, and that 00001 + 00100 < 00110.

Therefore, if $k \geqslant 5$, then $f^*(k) \geqslant \phi(k)$ where $\phi(k)$ satisfies the recurrences

$$\phi(2m+1) = 2\phi(2m) \quad \text{and} \quad \phi(2m) = 2\phi(2m-1) - 1,$$

with initial condition $\phi(5) = 16$. The unique solution is $\phi(k) = \lceil \frac{23}{48} \cdot 2^k \rceil$ for $k \geqslant 5$. We have seen that $f^*(4) = 8$ and we leave it to the reader to check that $f^*(k) = 2^{k-1}$ for $k = 1, 2$ and 3. $\square$

Corollary 2.8. $g(k) \geqslant f(k) \geqslant f^*(k) \geqslant \frac{23}{48} \cdot 2^k$.

3. Increasing paths in K_{2^k} and $K_{2^k,2^k}$

Let $n = 2^k$ where k is a positive integer. Consider the following edge ordering of the complete graph K_n (complete bipartite graph $K_{n,n}$). Label the vertices of K_n (each half of $K_{n,n}$) with the vectors of $\mathrm{GF}(2)^k$ and label the edge joining x to y with the vector $x + y$. Order edges with the same label in some fixed but arbitrary way and order edges with different labels lexicographically. Let $F(k)$ $(G(k))$ denote this edge-ordered graph.

Given an increasing path in K_n, the (modulo 2) sum of any subsequence of consecutive edge labels is never zero. Conversely, if we fix a vertex ν in $F(k)$, then the f-sequence $(a_i)_{i=1}^t$ determines the increasing path $\nu, \nu + a_1, \ldots, \nu + a_1 + \cdots + a_t$, in $F(k)$. Therefore $f(k)$ is the maximum length of a simple increasing path in $F(k)$. Similarly the maximum length of a simple increasing path in $G(k)$ is $g(k)$. Let $A = (a_i)_{i=1}^{g(k)}$ be a g-sequence in $\mathrm{GF}(2)^k$ of maximal length $g(k)$, and let γ be a simple increasing path in $G(k)$ determined by A. We shall now describe how γ determines an increasing '*pseudo-path*' $\gamma(t)$ in $G(t)$ for $0 < t < k$.

Given a positive integer $t < k$ and a vector $\nu = (\nu_{k-1}, \ldots, \nu_0)$ in $\mathrm{GF}(2)^k$, define the *t-prefix* $\nu(t)$ of ν by $\nu(t) = (\nu_{k-1}, \ldots, \nu_{k-t})$. If i is a vector in $\mathrm{GF}(2)^t$ let x_i be the number of vectors in A with the t-prefix i. The graph $G(t)$ is obtained from $G(k)$ by identifying vertices and edges labelled by vectors in $\mathrm{GF}(2)^k$ with the same t-prefix. An edge connecting ν and w in γ determines an edge connecting $\nu(t)$ and $w(t)$ in $G(t)$. This edge is directed $\nu(t) \to w(t)$ when $x_{(\nu+w)(t)}$ is odd and $\nu(t) \to \leftarrow w(t)$ when $x_{(\nu+w)(t)}$ is even. (To traverse the edge $\nu(t) \to \leftarrow w(t)$ cross from $\nu(t)$ to $w(t)$ and then cross back from $w(t)$ to $\nu(t)$.) The directed edges determined by the edges of γ in this way form the increasing *pseudo-path* $\gamma(t)$ in $G(t)$. (We use the term *pseudo-path* because consecutive edges in $\gamma(t)$ share a vertex but the graph underlying $\gamma(t)$ is not in general a simple path). The pseudo-path $\gamma(t)$ is called the *the t-projection of* γ and it is said to be *full* if it contains the maximum number of edges, namely 2^t. The path $\gamma(t)$ is full if and only if every vector in $\mathrm{GF}(2)^t$ is the t-prefix of some edge label in γ. If $\gamma(t-i)$ contains at most $2^{t-i}-1$ edges then $\gamma(t)$ contains at most 2^t-2^i edges. It follows that if $\gamma(t)$ is full, then $\gamma(t-i)$ is also full.

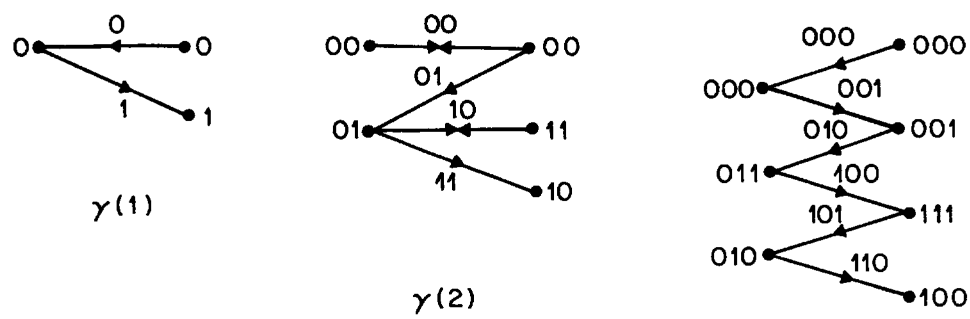
Example. $A=(000, 001, 010, 100, 101, 110)$

$t=1$,

i	0	1
x_i	3	3

; $t=2$,

i	00	01	10	11
x_i	2	1	2	1



Lemma 3.1. *Let d and t be fixed positive integers. For any $k>t$, suppose that a full t-projection of a simple increasing path in $G(k)$ always contains a vertex of degree at least d. Then*

$$\lim_{k\to\infty}(2^{-k}g(k))\leqslant 2/d.$$

Proof. By (2.2) the function $(2^{-k}g(k))$ is a decreasing function of k. By (2.8) we have $2^{-k}g(k)\geqslant 23/48$ for all k. Therefore $2^{-k}g(k)\downarrow c$ where c is a non-zero constant. We set $2^{-k}g(k)=c+\varepsilon(k)$ where $\varepsilon(k)\downarrow 0$.

Let γ be a simple increasing path in $G(k)$ $(k>t)$ of length $g(k)$. An edge in $\gamma(t)$ represents at most $g(k-t)$ edges in γ. Therefore, if $\gamma(t)$ is not full, then $g(k)\leqslant(2^t-1)g(k-t)$, or equivalently,

$$2^k(c+\varepsilon(k))\leqslant(2^t-1)2^{k-t}(c+\varepsilon(k-t)). \tag{3.2}$$

Letting $k\to\infty$ in (3.2) we obtain $c=0$ which is impossible. It follows that if k is sufficiently large, then $\gamma(t)$ is full. In that case it contains a vertex ν of degree at least d. A vertex in $\gamma(t)$ represents at most 2^{k-t} vertices in γ. Since γ is a simple path the d edges of $\gamma(t)$ that are joined to ν can only represent $2\times 2^{k-t}$ edges in γ. Hence

$$2^k(c+\varepsilon(k))\leqslant 2\cdot 2^{k-t}+(2^t-d)(2^{k-t}(c+\varepsilon(k-t))). \tag{3.3}$$

Letting $k\to\infty$ in (3.3) we obtain $c\leqslant 2/d$ as required. $\square$

Theorem 3.4. $\lim_{k\to\infty}(2^{-k}g(k))\leq\frac{1}{2}$.

Proof. It is sufficient to show that the hypothesis of Lemma 3.1 is satisfied with $d=4$ and $t=8$. The argument is rather lengthy and it is given in the appendix. □

Corollary 3.5. *Given any integer k,*

$$\alpha(2^k)\leq f(k)\leq g(k)\leq(\tfrac{1}{2}+o(1))2^k.$$

4. Increasing paths in K_n and $K_{n,n}$

Let n be some positive integer. Suppose that for all $m<n$ we have labelled the vertices of each half of $K_{m,m}$ with the numbers $1,\ldots,m$ and we have ordered the edges so that any pair of edges of the form (x,y), (y,x) are neighbors in the ordering. This edge ordering, $\mathbf{G}(m)$ say, of $K_{m,m}$ induces an edge ordering $\mathbf{F}(m)$ of K_m. Define an edge ordering $\mathbf{G}(n)$ of $K_{n,n}$ as follows. Label the vertices of each half of $K_{n,n}$ with different ordered pairs (x,y) where $x\in GF(2)$ and $1\leq y\leq\lceil\frac{1}{2}n\rceil=n'$. The edge from (x,y) to (x',y') is assigned the label $(x+x',\{y,y'\})$. Edges with the same label are ordered so that pairs of edges of the form $((x,y),(x',y'))$, $((x',y'),(x,y))$ are neighbors. Edges with different labels $(x_1+x_1',\{y_1,y_1'\})$, $(x_2+x_2',\{y_2,y_2'\})$ say, are ordered according to the first component and, if first components are equal, according to the ordering of edges (y_1,y_1'), (y_2,y_2') in $\mathbf{G}(n')$. As above, this edge ordering induces an edge ordering $\mathbf{F}(n)$ of K_n. Let $r(n)$ and $s(n)$ denote the maximum length of a simple increasing path in $\mathbf{F}(n)$ and $\mathbf{G}(n)$ respectively. It is easily seen that

$$r(n)\leq r(\lceil\tfrac{1}{2}n\rceil)+s(\lceil\tfrac{1}{2}n\rceil), \tag{4.1}$$

and

$$s(n)\leq 2s(\lceil\tfrac{1}{2}n\rceil). \tag{4.2}$$

Define $h(n)$ by $h(n)=2h(\lceil\frac{1}{2}n\rceil)$ and by setting $h(1)=1$. The function $(h(n)+\log(n))/n$ is decreasing for $n\geq 8$. Therefore $s(n)+\log(n)=(1+o(1))cn$ for some constant c. Let γ be an increasing path in $\mathbf{G}(n)$ of length $s(n)$. Again we can consider the t-projection of γ for $t\leq 8$. The method used to prove (3.1) and (3.4) allows us to conclude that $\gamma(8)$ is full and it contains a vertex of degree at least 4. Hence

$$s(n)\leq 2(\lceil n/256\rceil)+252s(\lceil n/256\rceil). \tag{4.3}$$

Let $n'=\lceil n/256\rceil$ and let $s(m)=(c+\varepsilon(m))m$ where $\varepsilon(m)\downarrow 0$ as $m\to\infty$. Equation (4.3) gives

$$(c+\varepsilon(n))\leq 2\cdot\frac{n}{256}+128+252(c+\varepsilon(n'))n'$$

and letting $n \to \infty$ we obtain $c \leqslant \frac{1}{2}$. We conclude that

$$r(n) \leqslant s(n) \leqslant (\tfrac{1}{2} + o(1))n$$

and so we have proved:

Theorem 4.4. $\alpha(n) \leqslant (\frac{1}{2} + o(1))n$.

Remarks. (1) The lower bound on $\alpha(n)$ obtained by Graham and Kleitman in [3] is about $\sqrt{n}$, and this would seem far from satisfactory. It does not appear unreasonable to ask if

$$\alpha(n) = (\tfrac{1}{2} + o(1))n.$$

(2) Is is true that $f(k) = g(k)$? It would also be of interest to know if either $f(k)$ or $g(k)$ is equal to $(1 + o(1))2^{k-1}$.

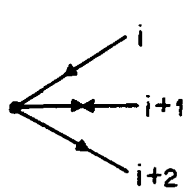
5. Appendix

This appendix contains a proof of Theorem 3.4. Recall that it is sufficient to show that the hypothesis of Lemma 3.1 is satisfied with $d = 4$ and $t = 8$. Suppose by way of contradiction that there exists a simple increasing path γ in $G(k)$, $k > 8$, such that $\gamma(8)$ is full and every vertex in $\gamma(8)$ has degree less than 4.

If $t < 8$, then $\gamma(t)$ is also full. A vertex $\nu(t)$ in $\gamma(t)$ splits into the two vertices $\nu(t)0$ and $\nu(t)1$ of $G(t+1)$. An edge labelled i in $\gamma(t)$ splits into two edges labelled $i0$ and $i1$ in $\gamma(t+1)$. If ν is a vertex of degree s in $\gamma(t)$, then by the pigeonhole principle one of the vertices $\nu 0$, $\nu 1$ in $\gamma(t+1)$ has degree $\geqslant s$. Hence the maximum degree in $\gamma(t)$ is less than 4.

Let O denote an odd positive integer and let E denote an even positive integer. If i is an edge label in $\gamma(t)$ and if x_i is odd, then x_i splits as $(x_{i0}, x_{i1}) = (O, E)$ or (E, O) in $\gamma(t+1)$. If x_i is even, then x_i splits as $(x_{i0}, x_{i1}) = (O, O)$ or (E, E). Henceforth we shall only label edges to indicate the order in which they are taken.

Claim. *There is a pseudo-path of the form*



(5.1)

in $\gamma(2)$ *or* $\gamma(3)$.

Proof. Suppose that the claim is false. There are 4 edges in $\gamma(2)$ and 7 possible edge sequences, namely (E, O, O, E), (E, O, E, E), (E, O, O, O), (E, E, O, E), (E, E, O, O),(O, O, O, O), and (O, O, O, E). These sequences determine the possible edge sequences for $\gamma(3)$, but none of the new sequences satisfies the initial hypotheses. For example, to avoid the forbidden pseudo-path (5.1), the edge sequence (E, O, O, E) must split as $(O, O, O, E, E, O, \ldots)$, $(E, E, O, E, E, O, \ldots)$, or $(E, E, E, O, \ldots)$ and in each case there is a vertex of degree 4. □

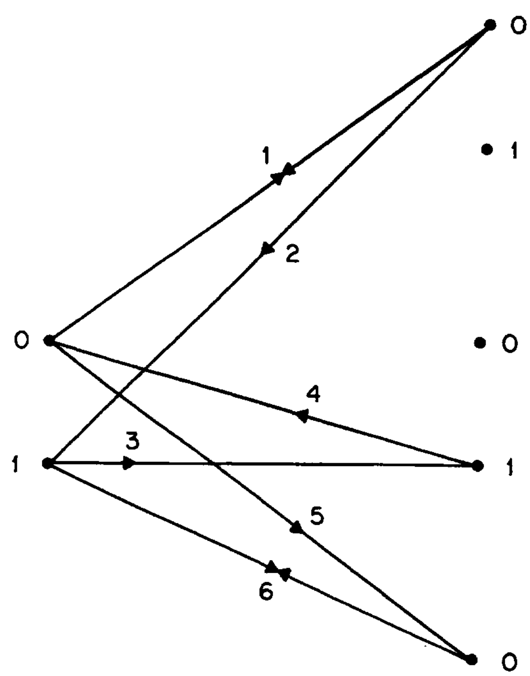
The pseudo-path of the form (5.1) contained in $\gamma(t)$ ($t=2$ or 3) determines a pseudo-path $\alpha(t+1)$ contained in $\gamma(t+1)$. It is easily seen that $\alpha(t+1)$ is one of the 2 configurations given below. (Note that an edge sequence $(\ldots, E, E, \ldots)$ is not allowed because it implies the existence of a vertex of degree at least 4.)

Case A.

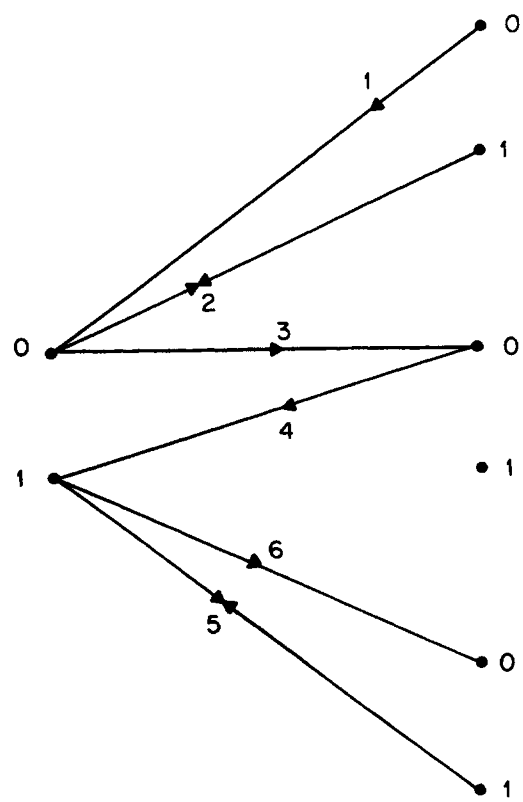
i	1	2	3	4	5	6
x_i	E	O	O	O	O	E

Case B.

i	1	2	3	4	5	6
x_i	O	E	O	O	E	O



(5.2)

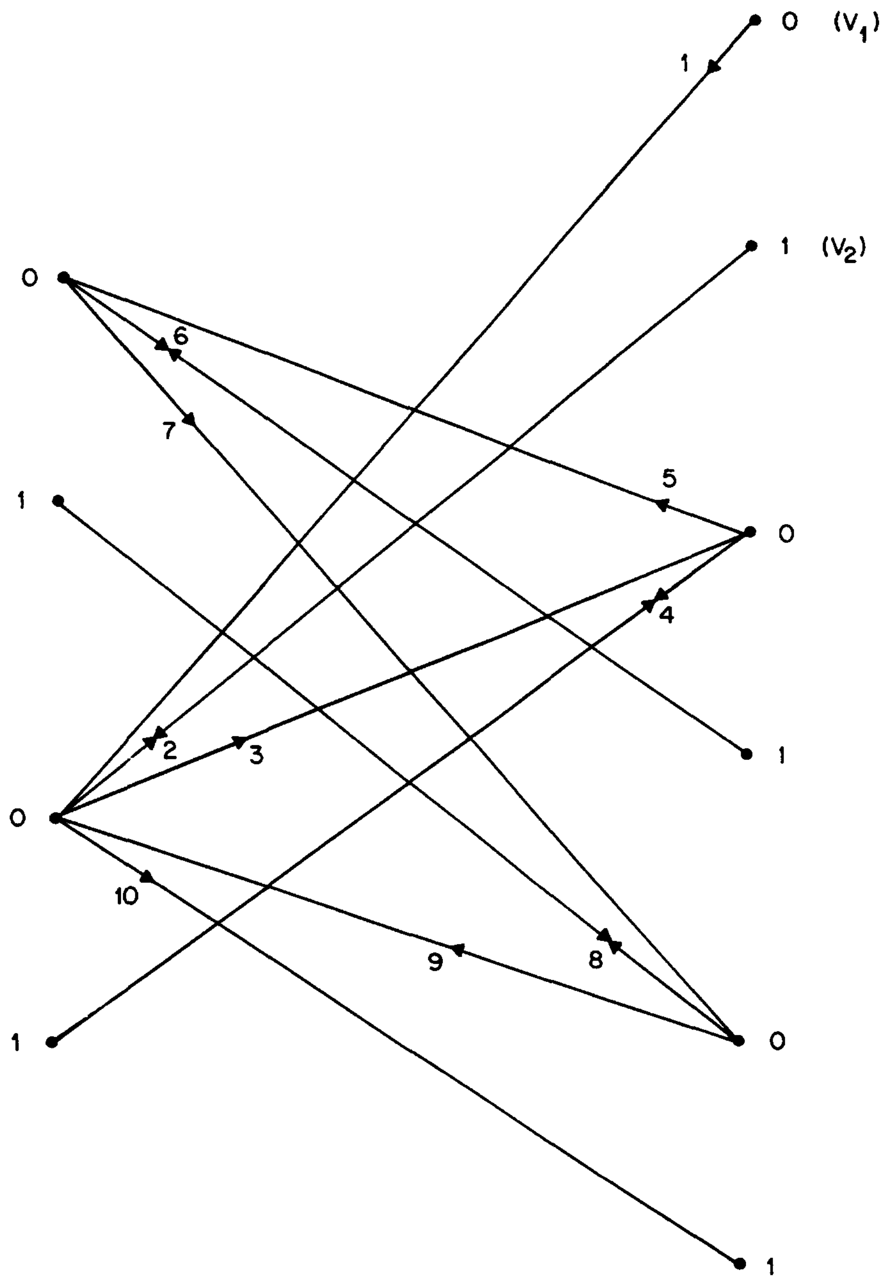


(5.3)

We begin with Case A. We consider the pseudo-path in $\gamma(t+2)$ that is determined by the edges 2, 3, 4, 5, and 6 of (5.2). If the odd edge 2 splits as

(O, E) then there is only one way to complete the new edge sequence and that is given below.

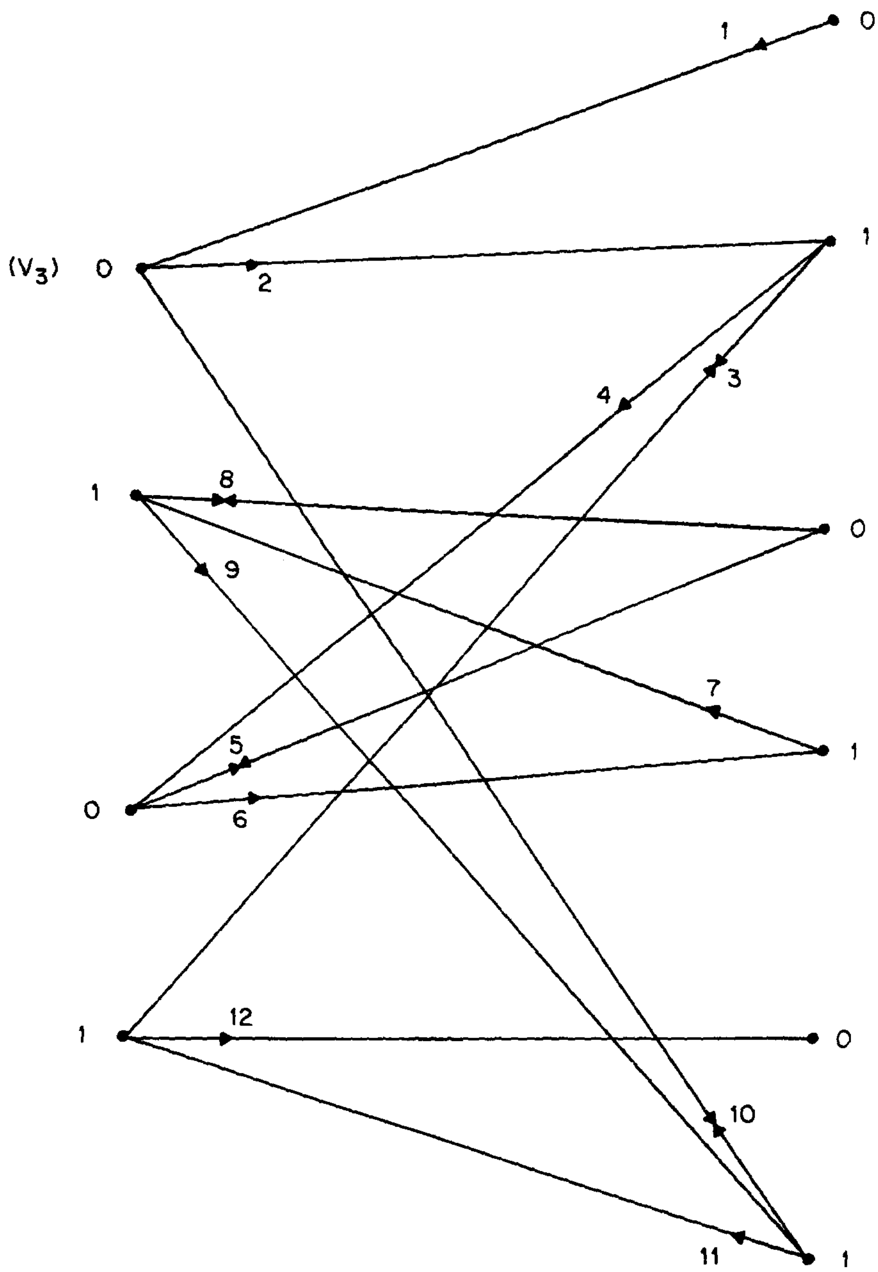
i	1	2	3	4	5	6	7	8	9	10
x_i	O	E	O	E	O	E	O	E	O	O



(5.4)

However (5.4) contradicts the nonexistence of a vertex of degree $\geqslant 4$. Note that if we had chosen ν_2 as the start of the path instead of ν_1 then we would have an isomorphic path and in particular we would still have a contradiction.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	O	O	E	O	E	O	O	E	O	E	O	O



(5.5)

We conclude that edges 1 and 2 in (5.2) split as (O, O, E, O). If edges 3 and 4 of (5.2) split as (O, E, O, E) then the vertex v_3 in (5.5) above has degree 4. Therefore the new edge sequence begins (O, O, E, O, E, O) and the only way to complete it is shown in (5.5).

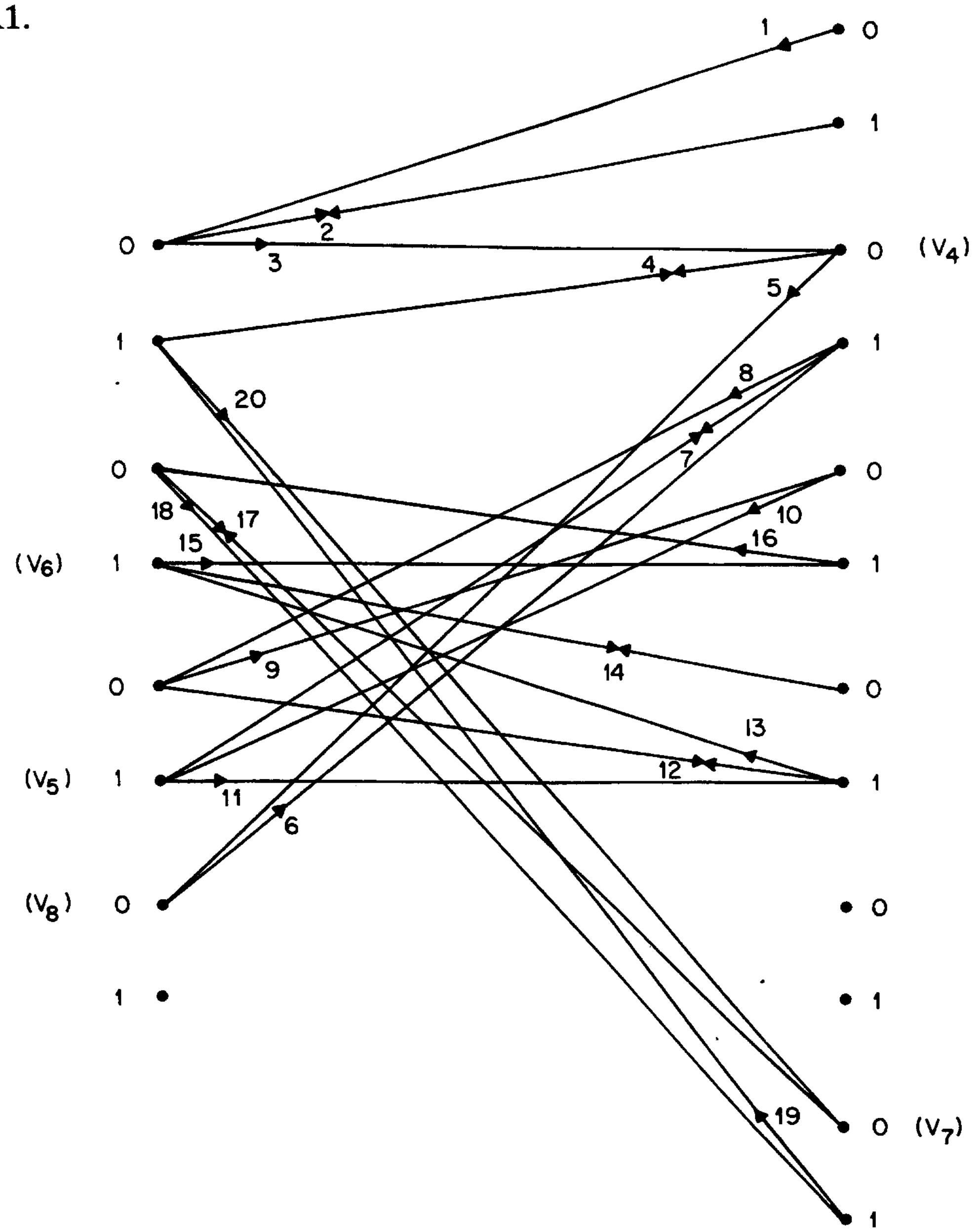
Finally we consider the pseudo-path in $\gamma(t+3)$ determined by (5.5). There are 3 cases.

Case A1. *Edges* 1 *and* 2 *of* (5.5) *split as* (O, E, O, E).

We claim that the new edge sequence begins

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	O	E	O	E	O	O	E	O	O	O	O	E	O	E	O	O	E	O	O	O

If $(7,8)=(O,E)$, then ν_4 in (5.6) has degree $\geqslant 4$. If $(11,12)=(E,O)$, then ν_5 has degree $\geqslant 4$. If $(17,18)=(O,E)$, then ν_6 has degree $\geqslant 4$. If $(21,22)=(E,O)$, then ν_7 has degree $\geqslant 4$ and if $(21,22)=(O,E)$, then ν_8 has degree $\geqslant 4$. This eliminates Case A1.



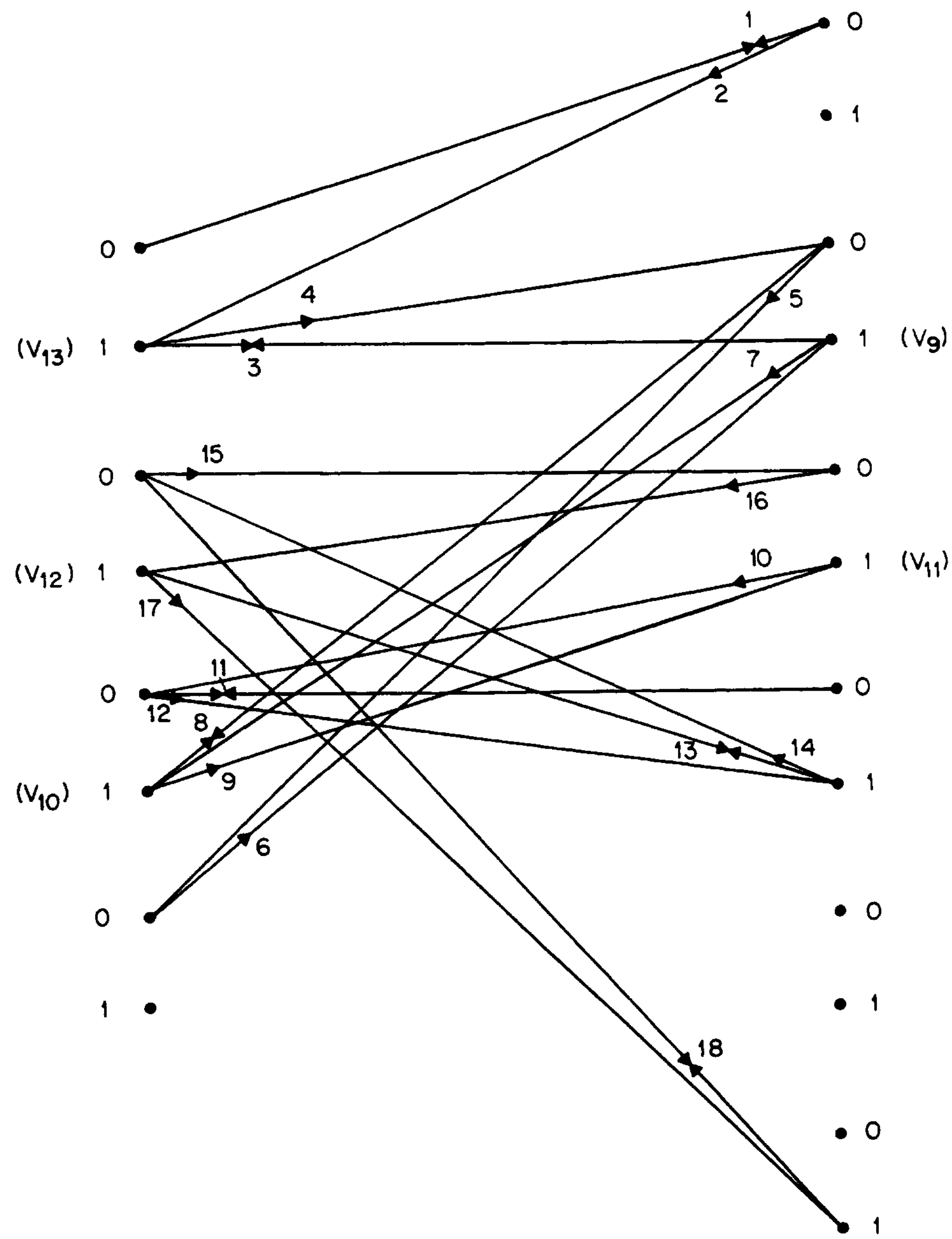
(5.6)

Case A2. *Edges* 1 *and* 2 *of* (5.5) *split as* (E,O,E,O).

We claim that the new edge sequence begins

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	E	O	E	O	O	O	O	E	O	O	E	O	E	O	O	O	O	E	O	O

If $(7, 8) = (E, O)$, then ν_9 in (5.7) has degree $\geqslant 4$. If $(11, 12) = (O, E)$, then ν_{10} has degree $\geqslant 4$. If $(13, 14) = (O, E)$, then $(13, 14, 15, 16) = (O, E, O, O)$ and ν_{11} has degree $\geqslant 4$. If $(17, 18) = (E, O)$, then ν_{12} has degree $\geqslant 4$. Now $(19, 20) = (O, O)$ and ν_{13} has degree $\geqslant 5$. This eliminates Case A2.

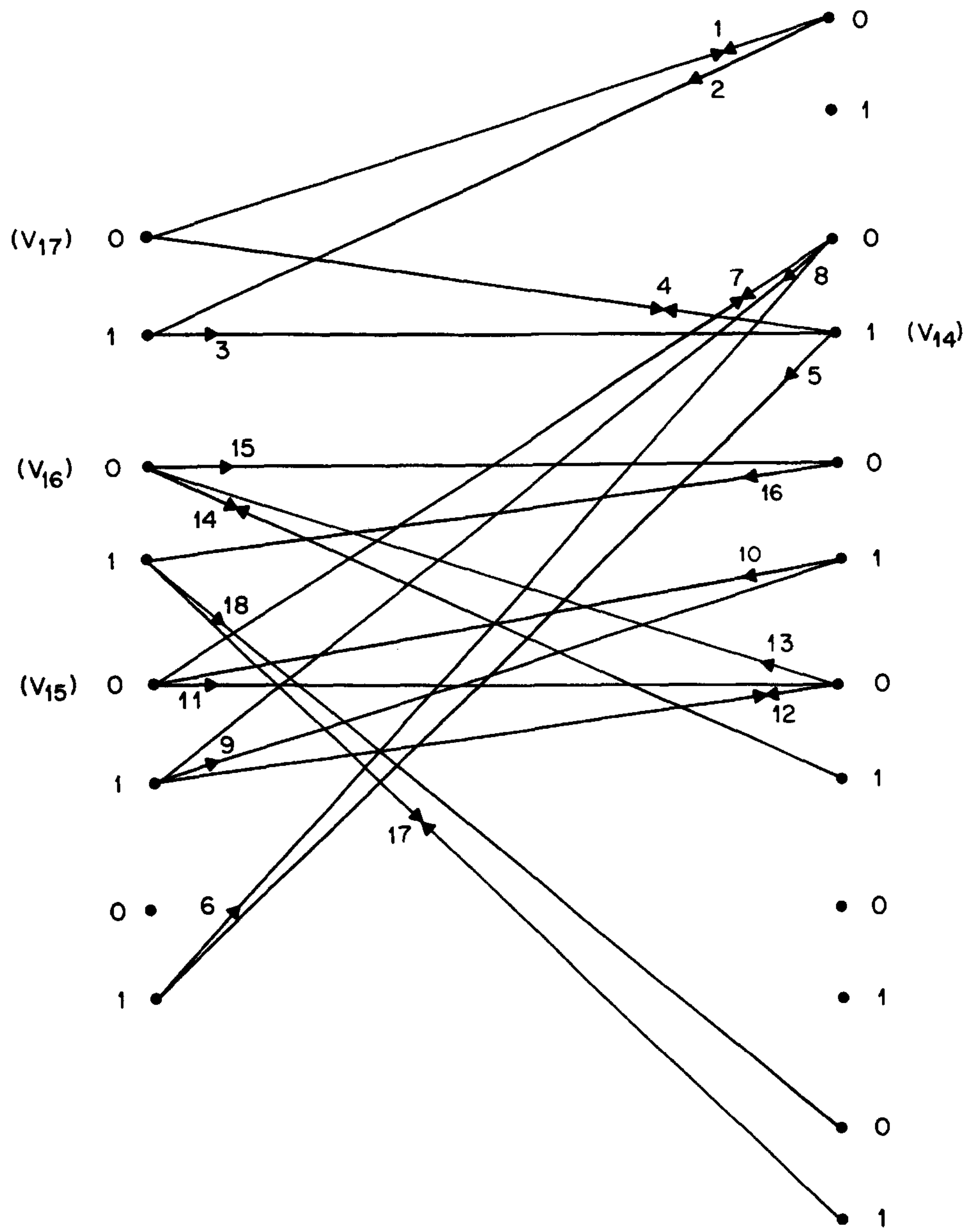


(5.7)

Case A3. *Edges* 1 *and* 2 *of* (5.5) *split as* (E, O, O, E).
We claim that the new edge sequence begins

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	E	O	O	E	O	O	E	O	O	O	O	E	O	E	O	O				

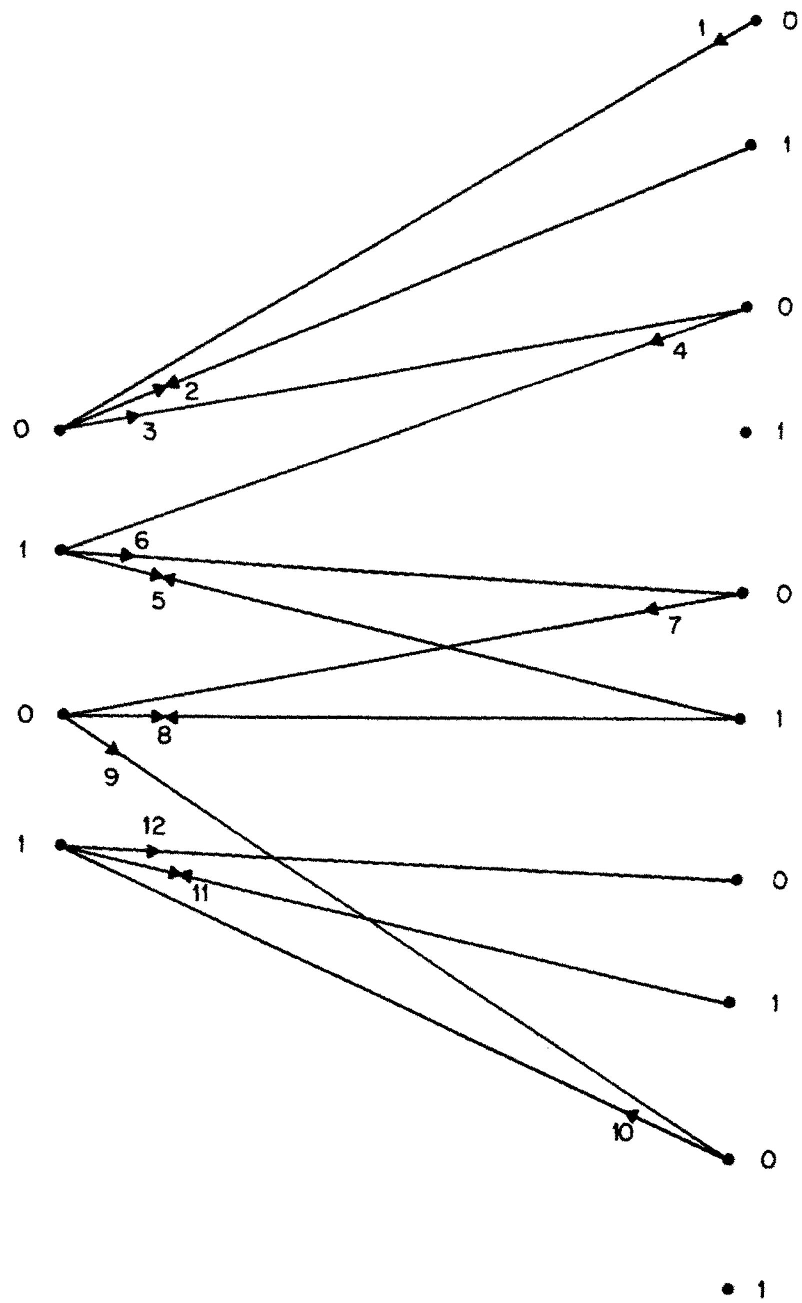
If $(7,8)=(O,E)$, then ν_{14} in (5.8) has degree$\geqslant 4$. If $(11,12)=(E,O)$, then ν_{15} has degree$\geqslant 4$. If $(17,18)=(O,E)$, then ν_{16} has degree$\geqslant 4$. Now $(19,20)=(O,O)$ and ν_{17} has degree$\geqslant 4$. This eliminates Case A3.



(5.8)

We have shown that if $\gamma(i)$ contains the pseudo-path (5.2), then $\gamma(i+2)$ contains a vertex of degree at least 4. We shall now eliminate Case B. We consider the pseudo-path in $\gamma(t+2)$ determined by (5.3). Since the new pseudo-path cannot contain (5.2) the new edge sequence must be

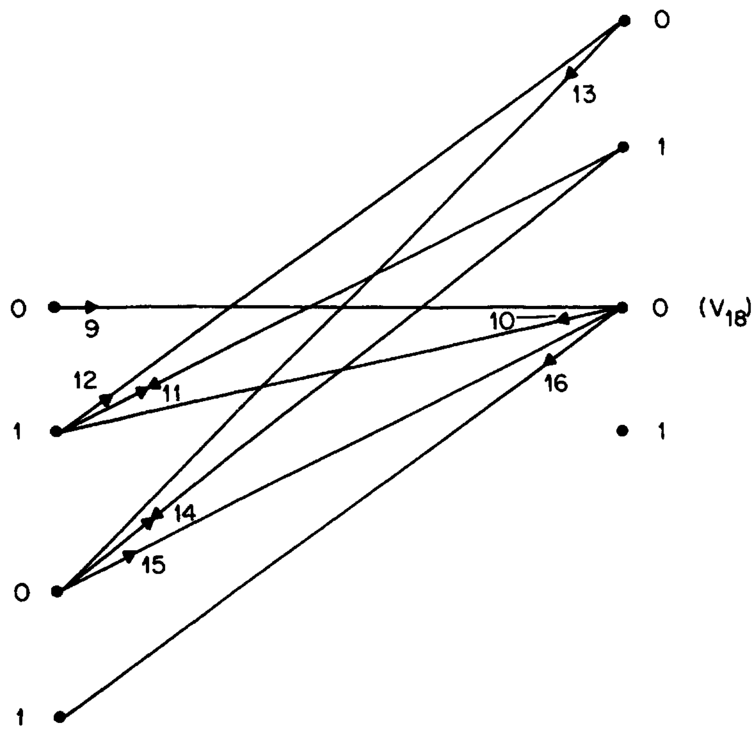
i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	O	O	E	O	E	O	O	E	O	E	O	O



(5.9)

Finally we consider the pseudo-path in $\gamma(t+3)$ determined by (5.9). Since the new pseudo-path cannot contain (5.2), the new edge sequence is 4 consecutive copies of (O, E, O, O, E, O). The new pseudo-path begins as in (5.9) and the edges $5, 6, 7$, and 8 of (5.9) split as in (5.10).

However the vertex ν_{18} now has degree at least 4. This is impossible and so the proof is complete. □



(5.10)

References

[1] B. Alspach, K. Heinrich, and R. L. Graham, Personal communication.

[2] V. Chvátal and J. Komlós, Some combinatorial theorems on monotonicity, Canad. Math. Bull. 14 (1971).

[3] R.L. Graham and D.J. Kleitman, Increasing paths in edge-ordered graphs, Per. Math. Hung. 3 (1973) 141–148.