



# COMBINATORIAL GAMES

# Game 1

Start with  $n$  chips. Players A,B alternately take 1,2,3 or 4 chips until there are none left. The winner is the person who takes the last chip:

Example

	A	B	A	B	A	
$n = 10$	3	2	4	1		B wins
$n = 11$	1	2	3	4	1	A wins

What is the optimal strategy for playing this game?

# Game 2

Chip placed at point  $(m, n)$ . Players can move chip to  $(m', n)$  or  $(m, n')$  where  $0 \leq m' < m$  and  $0 \leq n' < n$ . The player who makes the last move and puts the chip onto  $(0, 0)$  wins.

What is the optimal strategy for this game?

**Game 2a** Chip placed at point  $(m, n)$ . Players can move chip to  $(m', n)$  or  $(m, n')$  or to  $(m - a, n - a)$  where  $0 \leq m' < m$  and  $0 \leq n' < n$  and  $0 \leq a \leq \min\{m, n\}$ . The player who makes the last move and puts the chip onto  $(0, 0)$  wins.

What is the optimal strategy for this game?

# Game 3

$W$  is a set of words. A and B alternately remove words  $w_1, w_2, \dots$ , from  $W$ . The rule is that the first letter of  $w_{i+1}$  must be the same as the last letter of  $w_i$ . The player who makes the last legal move wins.

## Example

$W = \{ \textit{England, France, Germany, Russia, Bulgaria, \dots} \}$

What is the optimal strategy for this game?

# Abstraction

Represent each position of the game by a vertex of a digraph

$$D = (X, A).$$

$(x, y)$  is an arc of  $D$  iff one can move from position  $x$  to position  $y$ .

We assume that the digraph is finite and that it is **acyclic** i.e. there are no directed cycles.

The game starts with a token on vertex  $x_0$  say, and players alternately move the token to  $x_1, x_2, \dots$ , where  $x_{i+1} \in N^+(x_i)$ , the set of out-neighbours of  $x_i$ . The game ends when the token is on a **sink** i.e. a vertex of out-degree zero. The last player to move is the winner.

# Abstraction

Example 1:  $V(D) = \{0, 1, \dots, n\}$  and  $(x, y) \in A$  iff  $x - y \in \{1, 2, 3, 4\}$ .

Example 2:  $V(D) = \{0, 1, \dots, m\} \times \{0, 1, \dots, n\}$  and  $(x, y) \in N^+((x', y'))$  iff  $x = x'$  and  $y > y'$  or  $x > x'$  and  $y = y'$ .

Example 2a:  $V(D) = \{0, 1, \dots, m\} \times \{0, 1, \dots, n\}$  and  $(x, y) \in N^+((x', y'))$  iff  $x = x'$  and  $y > y'$  or  $x > x'$  and  $y = y'$  or  $x - x' = y - y' > 0$ .

Example 3:  $V(D) = \{(W', w) : W' \subseteq W \setminus \{w\}\}$ .  $w$  is the last word used and  $W'$  is the remaining set of unused words.  $(X', w') \in N^+((X, w))$  iff  $w' \in X$  and  $w'$  begins with the last letter of  $w$ . Also, there is an arc from  $(W, \cdot)$  to  $(W \setminus \{w\}, w)$  for all  $w$ , corresponding to the games start.

# Abstraction

We will first argue that such a game must eventually end.

A **topological numbering** of digraph  $D = (X, A)$  is a map  $f : X \rightarrow [n]$ ,  $n = |X|$  which satisfies  $(x, y) \in A$  implies  $f(x) < f(y)$ .

## Theorem

*A finite digraph  $D = (X, A)$  is acyclic iff it admits at least one topological numbering.*

**Proof**     Suppose first that  $D$  has a topological numbering. We show that it is acyclic.

Suppose that  $C = (x_1, x_2, \dots, x_k, x_1)$  is a directed cycle. Then  $f(x_1) < f(x_2) < \dots < f(x_k) < f(x_1)$ , contradiction.

# Abstraction

Suppose now that  $D$  is acyclic. We first argue that  $D$  has at least one sink.

Thus let  $P = (x_1, x_2, \dots, x_k)$  be a longest simple path in  $D$ . We claim that  $x_k$  is a sink.

If  $D$  contains an arc  $(x_k, y)$  then either  $y = x_i$ ,  $1 \leq i \leq k - 1$  and this means that  $D$  contains the cycle  $(x_i, x_{i+1}, \dots, x_k, x_i)$ , contradiction or  $y \notin \{x_1, x_2, \dots, x_k\}$  and then  $(P, y)$  is a longer simple path than  $P$ , contradiction.



# Abstraction

We can now prove by induction on  $n$  that there is at least one topological numbering.

If  $n = 1$  and  $X = \{x\}$  then  $f(x) = 1$  defines a topological numbering.

Now assume that  $n > 1$ . Let  $z$  be a sink of  $D$  and define  $f(z) = n$ . The digraph  $D' = D - z$  is acyclic and by the induction hypothesis it admits a topological numbering,  $f : X \setminus \{z\} \rightarrow [n - 1]$ .

The function we have defined on  $X$  is a topological numbering. If  $(x, y) \in A$  then either  $x, y \neq z$  and then  $f(x) < f(y)$  by our assumption on  $f$ , or  $y = z$  and then  $f(x) < n = f(z)$  ( $x \neq z$  because  $z$  is a sink).



# Abstraction

The fact that  $D$  has a topological numbering implies that the game must end. Each move increases the  $f$  value of the current position by at least one and so after at most  $n$  moves a sink must be reached.

The positions of a game are partitioned into 2 sets:

- $P$ -positions: The next player cannot win. The **previous** player can win regardless of the current player's strategy.
- $N$ -positions: The **next** player has a strategy for winning the game.

Thus an  $N$ -position is a **winning** position for the next player and a  $P$ -position is a **losing** position for the next player.

The main problem is to determine  $N$  and  $P$  and what the strategy is for winning from an  $N$ -position.

Let the vertices of  $D$  be  $x_1, x_2, \dots, x_n$ , in topological order.

## Labelling procedure

- 1  $i \leftarrow n$ , Label  $x_n$  with  $P$ .  $N \leftarrow \emptyset$ ,  $P \leftarrow \emptyset$ .
- 2  $i \leftarrow i - 1$ . If  $i = 0$  STOP.
- 3 Label  $x_i$  with  $N$ , if  $N^+(x_i) \cap P \neq \emptyset$ .
- 4 Label  $x_i$  with  $P$ , if  $N^+(x_i) \subseteq N$ .
- 5 goto 2.

The partition  $N, P$  satisfies

$$x \in N \text{ iff } N^+(x) \cap P \neq \emptyset$$

To play from  $x \in N$ , move to  $y \in N^+(x) \cap P$ .

# Abstraction

In Game 1,  $P = \{5k : k \geq 0\}$ .

In Game 2,  $P = \{(x, x) : x \geq 0\}$ .

## Lemma

*The partition into  $N, P$  satisfying  $x \in N$  iff  $N^+(x) \cap P \neq \emptyset$  is unique.*

**Proof** If there were two partitions  $N_i, P_i, i = 1, 2$ , let  $x_i$  be the vertex of highest topological number which is not in  $(N_1 \cap N_2) \cup (P_1 \cap P_2)$ . Suppose that  $x_i \in N_1 \setminus N_2$ .

But then  $x_i \in N_1$  implies  $N^+(x_i) \cap P_1 \cap \{x_{i+1}, \dots, x_n\} \neq \emptyset$  and  $x_i \in P_2$  implies  $N^+(x_i) \cap P_2 \cap \{x_{i+1}, \dots, x_n\} = \emptyset$ .

But  $P_1 \cap \{x_{i+1}, \dots, x_n\} = P_2 \cap \{x_{i+1}, \dots, x_n\}$ .

# Sums of games

Suppose that we have  $p$  games  $G_1, G_2, \dots, G_p$  with digraphs  $D_i = (X_i, A_i)$ ,  $i = 1, 2, \dots, p$ .

The sum  $G_1 \oplus G_2 \oplus \dots \oplus G_p$  of these games is played as follows. A position is a vector

$(x_1, x_2, \dots, x_p) \in X = X_1 \times X_2 \times \dots \times X_p$ . To make a move, a player chooses  $i$  such that  $x_i$  is not a sink of  $D_i$  and then replaces  $x_i$  by  $y \in N_i^+(x_i)$ . The game ends when each  $x_i$  is a sink of  $D_i$  for  $i = 1, 2, \dots, n$ .

Knowing the partitions  $N_i, P_i$  for game  $i = 1, 2, \dots, p$  does not seem to be enough to determine how to play the sum of the games.

We need more information. This will be provided by the **Sprague-Grundy Numbering**

# Sums of games

## Example

**Nim** In a one pile game, we start with  $a \geq 0$  chips and while there is a positive number  $x$  of chips, a move consists of deleting  $y \leq x$  chips. In this game the  $N$ -positions are the positive integers and the unique  $P$ -position is 0.

In general, Nim consists of the sum of  $n$  single pile games starting with  $a_1, a_2, \dots, a_n > 0$ . A move consists of deleting some chips from a non-empty pile.

Example 2 is Nim with 2 piles.

## Sprague-Grundy (*SG*) Numbering

For  $S \subseteq \{0, 1, 2, \dots\}$  let

$$\text{mex}(S) = \min\{x \geq 0 : x \notin S\}.$$

Now given an acyclic digraph  $D = X, A$  with topological ordering  $x_1, x_2, \dots, x_n$  define  $g$  iteratively by

- 1  $i \leftarrow n, g(x_n) = 0.$
- 2  $i \leftarrow i - 1.$  If  $i = 0$  STOP.
- 3  $g(x_i) = \text{mex}(\{g(x) : x \in N^+(x_i)\}).$
- 4 goto 2.

## Lemma

$$x \in P \leftrightarrow g(x) = 0.$$

**Proof**      Because

$$x \in N \text{ iff } N^+(x) \cap P \neq \emptyset$$

all we have to show is that

$$g(x) > 0 \text{ iff } \exists y \in N^+(x) \text{ such that } g(y) = 0.$$

But this is immediate from  $g(x) = \text{mex}(\{g(y) : y \in N^+(x)\})$   $\square$



# Sums of games

Another one pile subtraction game.

- A player can remove any even number of chips, but not the whole pile.
- A player can remove the whole pile if it is odd.

The terminal positions are 0 or 2.

## Lemma

$g(0) = 0$ ,  $g(2k) = k - 1$  and  $g(2k - 1) = k$  for  $k \geq 1$ .

# Sums of games

**Proof** 0,2 are terminal positions and so  $g(0) = g(2) = 0$ .  
 $g(1) = 1$  because the only position one can move to from 1 is 0. We prove the remainder by induction on  $k$ .

Assume that  $k > 1$ .

$$\begin{aligned}g(2k) &= \text{mex}\{g(2k-2), g(2k-4), \dots, g(2)\} \\ &= \text{mex}\{k-2, k-3, \dots, 0\} \\ &= k-1.\end{aligned}$$

$$\begin{aligned}g(2k-1) &= \text{mex}\{g(2k-3), g(2k-5), \dots, g(1), g(0)\} \\ &= \text{mex}\{k-1, k-2, \dots, 0\} \\ &= k.\end{aligned}$$



# Sums of games

We now show how to compute the **SG** numbering for a sum of games.

For binary integers  $a = a_m a_{m-1} \cdots a_1 a_0$  and  $b = b_m b_{m-1} \cdots b_1 b_0$  we define  $a \oplus b = c_m c_{m-1} \cdots c_1 c_0$  by

$$c_i = \begin{cases} 1 & \text{if } a_i \neq b_i \\ 0 & \text{if } a_i = b_i \end{cases}$$

for  $i = 1, 2, \dots, m$ .

So  $11 \oplus 5 = 14$ .

# Sums of games

## Theorem

If  $g_i$  is the **SG** function for game  $G_i$ ,  $i = 1, 2, \dots, p$  then the **SG** function  $g$  for the sum of the games  $G = G_1 \oplus G_2 \oplus \dots \oplus G_p$  is defined by

$$g(x) = g_1(x_1) \oplus g_2(x_2) \oplus \dots \oplus g_p(x_p)$$

where  $x = (x_1, x_2, \dots, x_p)$ .

For example if in a game of Nim, the pile sizes are  $x_1, x_2, \dots, x_p$  then the **SG** value of the position is

$$x_1 \oplus x_2 \oplus \dots \oplus x_p$$

# Sums of games

**Proof** It is enough to show this for  $p = 2$  and then use induction on  $p$ .

Write  $G = H \oplus G_p$  where  $H = G_1 \oplus G_2 \oplus \cdots \oplus G_{p-1}$ . Let  $h$  be the SG numbering for  $H$ . Then, if  $y = (x_1, x_2, \dots, x_{p-1})$ ,

$$\begin{aligned}g(x) &= h(y) \oplus g_p(x_p) \quad \text{assuming theorem for } p = 2 \\ &= g_1(x_1) \oplus g_2(x_2) \oplus \cdots \oplus g_{p-1}(x_{p-1}) \oplus g_p(x_p)\end{aligned}$$

by induction.

It is enough now to show, for  $p = 2$ , that

- A1 If  $x \in X$  and  $g(x) = b > a$  then there exists  $x' \in N^+(x)$  such that  $g(x') = a$ .
- A2 If  $x \in X$  and  $g(x) = b$  and  $x' \in N^+(x)$  then  $g(x') \neq g(x)$ .
- A3 If  $x \in X$  and  $g(x) = 0$  and  $x' \in N^+(x)$  then  $g(x') \neq 0$

# Sums of games

A1. Write  $d = a \oplus b$ . Then

$$a = d \oplus b = d \oplus g_1(x_1) \oplus g_2(x_2). \quad (1)$$

Now suppose that we can show that either

$$(i) \ d \oplus g_1(x_1) < g_1(x_1) \text{ or } (ii) \ d \oplus g_2(x_2) < g_2(x_2) \text{ or both.} \quad (2)$$

Assume that (i) holds.

Then since  $g_1(x_1) = \text{mex}(N_1^+(x_1))$  there must exist  $x'_1 \in N_1^+(x_1)$  such that  $g_1(x'_1) = d \oplus g_1(x_1)$ .

Then from (1) we have

$$a = g_1(x'_1) \oplus g_2(x_2) = g(x'_1, x_2).$$

Furthermore,  $(x'_1, x_2) \in N^+(x)$  and so we will have verified A1.

# Sums of games

Let us verify (2).

Suppose that  $2^{k-1} \leq d < 2^k$ .

Then  $d$  has a 1 in position  $k$  and no higher.

Since  $d_k = a_k \oplus b_k$  and  $a < b$  we must have  $a_k = 0$  and  $b_k = 1$ .

So either (i)  $g_1(x_1)$  has a 1 in position  $k$  or (ii)  $g_2(x_2)$  has a 1 in position  $k$ . Assume (i).

But then  $d \oplus g_1(x_1) < g_1(x_1)$  since  $d$  “destroys” the  $k$ th bit of  $g_1(x_1)$  and does not change any higher bit.

# Sums of games

A2. Suppose without loss of generality that  $g(x'_1, x_2) = g(x_1, x_2)$  where  $x'_1 \in N^+(x_1)$ .

Then  $g_1(x'_1) \oplus g_2(x_2) = g_1(x_1) \oplus g_2(x_2)$  implies that  $g_1(x'_1) = g_1(x_1)$ , contradiction. □

A3. Suppose that  $g_1(x_1) \oplus g_2(x_2) = 0$  and  $g_1(x'_1) \oplus g_2(x_2) = 0$  where  $x'_1 \in N^+(x_1)$ .

Then  $g_1(x_1) = g_1(x'_1)$ , contradicting  $g_1(x_1) = \text{mex}\{g_1(x) : x \in N^+(x_1)\}$ .



# Sums of games

If we apply this theorem to the game of Nim then if the position  $x$  consists of piles of  $x_i$  chips for  $i = 1, 2, \dots, p$  then

$$g(x) = x_1 \oplus x_2 \oplus \dots \oplus x_p.$$

In our first example,  $g(x) = x \bmod 5$  and so for the sum of  $p$  such games we have

$$g(x_1, x_2, \dots, x_p) = (x_1 \bmod 5) \oplus (x_2 \bmod 5) \oplus \dots \oplus (x_p \bmod 5).$$

# A more complicated one pile game

Start with  $n$  chips. First player can remove up to  $n - 1$  chips.

In general, if the previous player took  $x$  chips, then the next player can take  $y \leq x$  chips.

Thus a games position can be represented by  $(n, x)$  where  $n$  is the current size of the pile and  $x$  is the maximum number of chips that can be removed in this round.

## Theorem

*Suppose that the position is  $(n, x)$  where  $n = m2^k$  and  $m$  is odd. Then,*

- (a) This is an  $N$ -position if  $x \geq 2^k$ .*
- (b) This is a  $P$ -position if  $m = 1$  and  $x < n$ .*

# A more complicated one pile game

**Proof** For a non-negative integer  $n = m2^k$ , let  $\text{ones}(n)$  denote the number of ones in the binary expansion of  $n$  and let  $k = \rho(n)$  determine the position of the right-most one in this expansion.

We claim that the following strategy is a win for the player in a position described in (a):

Remove  $y = 2^k$  chips.

Suppose this player is **A**.

If  $m = 1$  then  $x \geq n$  and **A** wins.

# A more complicated one pile game

Otherwise, after such a move the position is  $(n', y)$  where  $\rho(n') > \rho(n)$ .

Note first that  $\text{ones}(n') = \text{ones}(n) - 1 > 0$  and  $\rho(n') > k$ . **B** cannot remove more than  $2^k$  chips and so **B** cannot win at this point.

If **B** moves the position to  $(n'', x'')$  then  $\text{ones}(n'') > \text{ones}(n')$  and furthermore,  $x'' \geq 2^{\rho(n'')}$ , since  $x''$  must have a 1 in position  $\rho(n'')$ . ( $\rho(n'')$  is the least significant bit of  $x''$ .)

Thus, by induction, **A** is in an  $N$ -position and wins the game.

To prove (b), note that after the first move, the position satisfies the conditions of (a). □.

# A General Subtraction Game

Let us next consider a generalisation of this game.

There are 2 players **A** and **B** and **A** goes first.

We have a non-decreasing function  $f$  from  $\mathbf{N} \rightarrow \mathbf{N}$  where  $\mathbf{N} = \{1, 2, \dots\}$  which satisfies  $f(x) \geq x$ .

At the first move **A** takes any number less than  $h$  from the pile, where  $h$  is the size of the initial pile.

Then on a subsequent move, if a player takes  $x$  chips then the next player is constrained to take at most  $f(x)$  chips.

Thus the previous analysis was for the game with  $f(x) = x$ .

# A General Subtraction Game

There is a set  $\mathcal{H} = \{H_1 = 1 < H_2 < \dots\}$  of initial pile sizes for which the first player will lose, assuming that the second player plays optimally.

Also, if the initial pile size  $h \notin \mathcal{H}$  then the first player has a winning strategy. It will turn out that the sequence satisfies the recurrence:

$$H_{j+1} = H_j + H_\ell \text{ where } H_\ell = \min_{i \leq j} \{H_i \mid f(H_i) \geq H_j\}, \quad \text{for } j \geq 0.$$

If  $f(x) = x$  then  $H_j = 2^{j-1}$ .

We prove this inductively. It is true for  $j = 1$ .

$$\begin{aligned}H_{j+1} &= 2^{j-1} + \min_{i \leq j} \{2^{i-1} : 2^{i-1} \geq 2^{j-1}\} \\ &= 2^{j-1} + 2^{j-1} \\ &= 2^j.\end{aligned}$$

# A General Subtraction Game

If  $f(x) = 2x$  then  $\mathcal{H} = \{1, 2, 3, 5, 8, \dots\} = \{F_1, F_2, \dots\}$ , the Fibonacci sequence.

We prove this inductively. It is true for  $j = 1, 2$ .

$$\begin{aligned}H_{j+1} &= F_j + \min_{i \leq j} \{F_i : 2F_i \geq F_j\} \\ &= F_j + F_{j-1} \\ &= F_{j+1}.\end{aligned}$$

Recall that  $F_j = F_{j-1} + F_{j-2}$  and

$$2F_{j-2} < F_{j-1} + F_{j-2} = F_j.$$



# A General Subtraction Game

The key to the game is the following result.

## Theorem

Every positive integer  $n$  can be *uniquely* written as the sum

$$n = H_{j_1} + H_{j_2} + \cdots + H_{j_p}$$

where  $f(H_{j_i}) < H_{j_{i+1}}$  for  $1 \leq i < p$ .

One simple consequence of the uniqueness of the decomposition is that

$$H_k \neq H_{j_1} + H_{j_2} + \cdots + H_{j_p}$$

for all  $k$  and sequences  $j_1, j_2, \dots, j_p$  where  $f(H_{j_i}) < H_{j_{i+1}}$  for  $i = 1, 2, \dots, p - 1$ .

# A General Subtraction Game

It follows that the integers  $n$  can be given unique “binary” representations by representing  $n = H_{j_1} + H_{j_2} + \dots + H_{j_p}$  by the 0-1 string with a 1 in positions  $j_1, j_2, \dots, j_p$  and 0 everywhere else.

Let  $\rho_H(n) = p$  be the number of 1's in the representation.

We call this the  $H$ -representation of  $n$ . This then leads to the following

## Theorem

*Suppose that the start position is  $(n, *)$ . Then,*

- (a) This is an  $N$ -position if  $n \notin \mathcal{H} = \{H_1, H_2, \dots\}$ .*
- (b) This is a  $P$ -position if  $n \in \mathcal{H}$ .*

# A General Subtraction Game

(a) The winning strategy is to delete a number of chips equal to  $H_{j_1}$  where  $j_1$  is the index of the rightmost 1 in the  $H$ -representation of  $n = H_{j_p} + \cdots + H_{j_1}$ .

All we have to do is verify that this strategy is possible.

Note first that if **A** deletes  $H_{j_1}$  chips, then **B** cannot respond by deleting  $H_{j_2}$  chips, because  $H_{j_2} > f(H_{j_1})$ .

**B** is forced to delete  $x \leq f(H_{j_1}) < H_{j_2}$  chips.

If  $p = 2$  then  $\rho_H(n - H_{j_1} - x) \geq 1 = \rho_H(n - H_{j_1})$ .

# A General Subtraction Game

If  $p \geq 3$  and  $y = H_{j_2} - x = H_{k_q} + \cdots + H_{k_1}$  then the  $H$ -representation of  $n - H_{j_1} - x$  is

$$H_{j_p} + \cdots + H_{j_3} + H_{k_q} + \cdots + H_{k_1}.$$

Here we use the fact that  $f(H_{k_q}) \leq f(y) \leq f(H_{j_2}) < H_{j_3}$ .

And so in both cases  $\rho_H(n - H_{j_2} - x) \geq \rho_H(n - H_{j_1})$  it is only **A** that can reduce  $\rho_H$ .

# A General Subtraction Game

The next thing to check is that if **A** starts in  $(n, *)$  then **A** can always delete  $H_{j_1}$  chips i.e. the positions  $(m, x)$  that **A** will face satisfy  $f(x) \geq H_{k_1}$  where  $m = H_{k_1} + H_{k_2} + \dots + H_{k_q}$ .

We do this by induction on the number of plays in the game so far.

It is true in the first move and suppose that it is true for  $(m, x)$  and that **A** removes  $H_{k_1}$  and **B** removes  $y$  where  $y \leq \min\{m - H_{k_1}, f(H_{k_1})\} < H_{k_2}$ . Now if  $H_{k_2} - y = H_{\ell_r} + H_{\ell_{r-1}} + \dots + H_{\ell_1}$  then

$$\begin{aligned} m - H_{k_1} - y &= H_{k_q} + \dots + H_{k_3} + H_{k_2} - y \\ &= H_{k_q} + \dots + H_{k_3} + H_{\ell_r} + H_{\ell_{r-1}} + \dots + H_{\ell_1} \end{aligned}$$

and we need to argue that  $H_{\ell_1} \leq f(y)$ .

# A General Subtraction Game

But if  $f(y) < H_{\ell_1}$  then we have

$$\begin{aligned}H_{k_2} &= y + H_{\ell_1} + H_{\ell_2} + \cdots + H_{\ell_r} \\ &= H_{a_1} + \cdots + H_{a_s} + H_{\ell_1} + H_{\ell_2} + \cdots + H_{\ell_r}\end{aligned}$$

where  $f(H_{a_s}) \leq f(y) < H_{\ell_1}$ , which gives two distinct decompositions for  $H_{k_2}$ , contradiction.

Thus **A** can remove  $H_{\ell_1}$  in the next round, as required.

# A General Subtraction Game

(b) Assume that  $n = H_k$ . After **A** removes  $x$  chips we have

$$H_k - x = H_{j_1} + H_{j_2} + \cdots + H_{j_p}$$

chips left.

All we have to show is that **B** can now remove  $H_{j_1}$  chips i.e.  $H_{j_1} \leq f(x)$ .

But if this is not the case then we argue as above that  $H_k = H_{a_1} + \cdots + H_{a_s} + H_{j_1} + H_{j_2} + \cdots + H_{j_p}$ , where  $x = H_{a_1} + \cdots + H_{a_s}$  and  $f(H_{j_1}) \leq f(x) < H_{j_1}$ , which gives two distinct decompositions for  $H_k$ , contradiction.

# A General Subtraction Game

## Proof of the existence of a unique decomposition

We prove this by induction on  $n$ . If  $n = 1$  then  $n = H_1$  is the unique decomposition.

Going back to the defining recurrence we see that

$$H_{j+1} = H_j + H_\ell \leq 2H_j.$$

## Existence

Assume that any  $n < H_k$  can be represented as a sum of distinct  $H_{j_i}$ 's with  $f(H_{j_i}) < H_{j_{i+1}}$  and suppose that

$H_k \leq n < H_{k+1}$ .  $H_{k+1} \leq 2H_k$  implies that  $n - H_k < H_k$ .

It follows by induction that

$$n - H_k = H_{j_1} + \cdots + H_{j_p},$$

where  $f(H_{j_i}) < H_{j_{i+1}}$  for  $i = 1, 2, \dots, p - 1$ .



# A General Subtraction Game

Assume to the contrary that  $f(H_{j_p}) \geq H_k$ .

Then for some  $m \leq j_p$  we have

$$H_{k+1} = H_k + H_m \leq H_k + H_{j_p} \leq n,$$

contradicting the choice of  $n$ .

# A General Subtraction Game

## Uniqueness

We will first prove by induction on  $p$  that if  $f(H_{j_i}) < H_{j_{i+1}}$  for  $1 \leq i < p$  then

$$H_{j_1} + H_{j_2} + \cdots + H_{j_p} < H_{j_{p+1}}. \quad (3)$$

If  $p = 2$  then we are saying that if  $f(H_{j_1}) < H_{j_2}$  then  $H_{j_1} + H_{j_2} < H_{j_{2+1}}$ . But this follows directly from  $H_{j_{2+1}} = H_{j_2} + H_m$  where  $f(H_m) \geq H_{j_2}$  i.e.  $H_m > H_{j_1}$ .

So assume that (3) is true for  $p \geq 2$ . Now

$$H_{j_{p+1}+1} = H_{j_{p+1}} + H_m \text{ and } f(H_{j_p}) < H_{j_{p+1}}$$

implies that  $m \geq j_p + 1$ . Thus

$$\begin{aligned} H_{j_{p+1}+1} &\geq H_{j_{p+1}} + H_{j_p+1} \\ &> H_{j_{p+1}} + H_{j_p} + H_{j_{p-1}} + \cdots + H_{j_1} \end{aligned}$$

after applying induction to get the second inequality. This completes the induction for (3).

# A General Subtraction Game

Now assume by induction on  $k$  that  $n < H_k$  has a unique decomposition. This is true for  $k = 2$  and so now assume that  $k \geq 2$  and  $H_k \leq n < H_{k+1}$ . Consider a decomposition

$$n = H_{j_1} + H_{j_2} + \cdots + H_{j_p}.$$

It follows from (3) that  $j_p = k$ . Indeed,  $j_p \leq k$  since  $n < H_{k+1}$  and if  $j_p < k$  then  $H_{j_1} + H_{j_2} + \cdots + H_{j_p} < H_{j_p+1} \leq H_k$ , contradicting our choice of  $n$ . So  $H_k$  appears in every decomposition of  $n$ .

Now  $H_{k+1} \leq 2H_k$  and  $n < H_{k+1}$  implies  $n - H_k < H_k$  and so, by induction,  $n - H_k$  has a unique decomposition. But then if  $n$  had two distinct decompositions,  $H_k$  would appear in each, implying that  $n - H_k$  also had two distinct decompositions, contradiction.

Note that although we know the optimal strategy for this game, we do not know the Sprague-grundy numbers and so we do not immediately get a solution to multi-pile versions.

# Wythoff's Nim

This is Game 2a.

## Theorem

The set of  $P$ -positions is  $\mathcal{A} = ((a_i, b_i), i = 0, 1, 2, \dots)$  where  $a_i < b_i, i \neq 0$  can be generated as follows:  $a_0 = b_0 = 0$  and

- $a_i$  is the smallest integer not appearing in  $a_0, b_0, \dots, a_{i-1}, b_{i-1}$
- $b_i = a_i + i$ .

The sequence  $\mathcal{A}$  starts

0	0	1	2	3	5
4	7	6	10	8	13
9	15	11	18	12	20
14	23	16	26	17	28
19	31	21	34	22	36
24	39	25	41	27	44

# Wythoff's Nim

**Proof** We first prove that each positive integer appears exactly once either as  $a_i$  or  $b_i$ .

We cannot have  $a_i = a_j$  for  $i < j$  because  $a_j$  is the smallest integer that has not previously appeared. Similarly, we cannot have  $a_i < a_{i-1}$ , else  $a_{i-1}$  was too large.

Since  $b_i = a_i + i$  we see that both of the sequences  $a_0, a_1, \dots$ , and  $b_0, b_1, \dots$ , are monotone increasing.

Suppose then that  $x = a_i = b_j$ . Since  $a_i < b_i < b_j$  for  $i < j$ , we must have  $i > j$  here. But then  $a_i$  is not an integer that has not appeared before.

Thus each positive integer appears exactly once either as  $a_i$  or  $b_i$ .

# Wythoff's Nim

Now suppose that  $(a_i, b_i) \in \mathcal{A}$ . We consider the possible positions we can move to and check that we cannot move to  $\mathcal{A}$ :

①  $(a_i - x, b_i) = (a_j, b_j)$  where  $x > 0$ .

We must have  $j < i$  and  $b_j = b_i$ . Not possible.

②  $(a_i, b_i - x) = (a_j, b_j)$  where  $x > 0$ .

We must have  $j < i$  and  $a_j = a_i$ . Not possible.

③  $(a_i - x, b_i - x) = (a_j, b_j)$  where  $x > 0$ .

We must have  $j < i$  and  $i = b_i - a_i = b_j - a_j = j$ . Not possible.

# Wythoff's Nim

Now suppose that  $(c, d) \notin \mathcal{A}$ ,  $c, d$ . We see that we can move to a pair in  $\mathcal{A}$ .

①  $c = a_i$  and  $d > b_i$ .

We can move to  $(a_i, b_i)$  by removing  $d - b_i$  from the  $d$  pile.

②  $c = a_i$  and  $d < b_i$ .

Let  $j = d - c$ . We can move to  $(a_j, b_j)$  by deleting  $c - a_j = d - b_j$  from each pile.

③  $d = b_i$  and  $c > a_i$ .

We can move to  $(a_i, b_i)$  by removing  $c - a_i$  from the  $c$  pile.

④  $d = b_i$  and  $c < a_i$  and we are not in Case 1 (with  $i$  replaced by  $i'$ ).

Thus,  $c = b_j$  for some  $j < i$ . We can move to  $(a_j, b_j)$  by removing  $d - a_j$  from the  $d$  pile.

We have therefore verified that the sequence  $\mathcal{A}$  does indeed define the set of  $P$  positions.

# Wythoff's Nim

We can give the following description of the sequence  $\mathcal{A}$ .

## Theorem

$$a_k = \lfloor \frac{k}{2}(1 + \sqrt{5}) \rfloor \text{ and } b_k = \lfloor \frac{k}{2}(3 + \sqrt{5}) \rfloor$$

for  $k = 0, 1, 2, \dots$

**Proof** It will be enough to show that each non-negative integer appears exactly once in the sequence  $(x_k, y_k) = (\lfloor \frac{k}{2}(1 + \sqrt{5}) \rfloor, \lfloor \frac{k}{2}(3 + \sqrt{5}) \rfloor)$  (\*).

Given (\*) we assume inductively that  $(a_i, b_i) = (x_i, y_i)$  for  $0 \leq i \leq k$ . This is true for  $k = 0$ .

Using (\*) we see that  $a_{k+1}$  appears in some pair  $x_j, y_j$ . We must have  $j > k$  else  $a_{k+1}$  will appear in  $a_0, \dots, b_k$ .



# Wythoff's Nim

Now  $x_{k+1}$  is the smallest integer that does not appear in  $(x_0, \dots, y_k) = (a_0, \dots, b_k)$  and so  $x_{k+1} = a_{k+1}$  and then  $y_{k+1} = x_{k+1} + k = b_{k+1}$ , completing the induction.

# Wythoff's Nim

## Proof of (\*)

Fix an integer  $n$  and write

$$\alpha = \frac{1}{2}p(1 + \sqrt{5}) - n \quad (4)$$

$$\beta = \frac{1}{2}q(3 + \sqrt{5}) - n \quad (5)$$

where  $p, q$  are integers and

$$0 < \alpha < \frac{1}{2}p(1 + \sqrt{5}) \quad (6)$$

$$0 < \beta < \frac{1}{2}q(3 + \sqrt{5}) \quad (7)$$

# Wythoff's Nim

Multiply (9) by  $\frac{1}{2}(-1 + \sqrt{5})$  and (10) by  $\frac{1}{2}(3 - \sqrt{5})$  and add to get

$$\frac{1}{2}\alpha(-1 + \sqrt{5}) + \frac{1}{2}\beta(3 - \sqrt{5}) = p + q - n = \textit{integer}.$$

Multiply (6) by  $\frac{1}{2}(-1 + \sqrt{5})$  and (7) by  $\frac{1}{2}(3 - \sqrt{5})$  and add to get

$$0 < \frac{1}{2}\alpha(-1 + \sqrt{5}) + \frac{1}{2}\beta(3 - \sqrt{5}) < 2.$$

We see therefore that

$$\frac{1}{2}\alpha(-1 + \sqrt{5}) + \frac{1}{2}\beta(3 - \sqrt{5}) = p + q - n = 1. \quad (8)$$

Although  $\alpha = \beta = 1$  satisfies (8) this can be rejected by observing that (9) would then imply that  $n + 1 = p(1 + \sqrt{5})$ .

# Wythoff's Nim

Thus either (i)  $\alpha < 1, \beta > 1$  or (ii)  $\alpha > 1, \beta < 1$ .

In case (i) we have from (9) that  $n = \lfloor p(1 + \sqrt{5}) \rfloor$ , while in case (ii) we have from (10) that  $n = \lfloor q(3 + \sqrt{5}) \rfloor$

This proves that  $n$  appears among the  $x_k, y_k$ . We now argue that the  $x_k, y_k$  are distinct.

In Case (i) we can that since  $\beta > 1$  is as small as possible,  $n \neq y_k$  for every  $k$ . In Case (ii) we see that  $n \neq x_k$  for every  $k$ .

So if an  $n$  appears twice, then we would have (a)  $x_k = x_\ell$  or (b)  $y_k = y_\ell$  for some  $k > \ell$ .

But (a) implies  $0 = x_k - x_\ell = \frac{1}{2}(k - \ell)(1 + \sqrt{5}) - \eta$  where  $|\eta| < 1$ , a contradiction. We rule out (b) in the same way.

# Geography

Start with a chip sitting on a vertex  $v$  of a graph or digraph  $G$ . A move consists of moving the chip to a neighbouring vertex.

In edge geography, moving the chip from  $x$  to  $y$  deletes the edge  $(x, y)$ . In vertex geography, moving the chip from  $x$  to  $y$  deletes the vertex  $x$ .

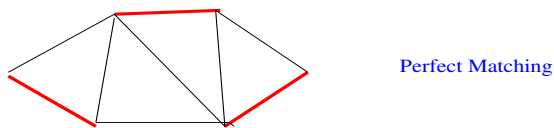
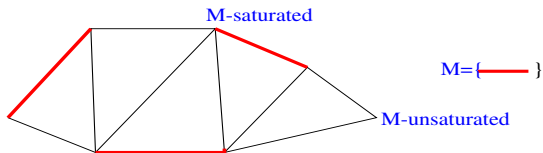
The problem is given a position  $(G, v)$ , to determine whether this is a  $P$  or  $N$  position.

**Complexity** Both edge and vertex geography are Pspace-hard on digraphs. Edge geography is Pspace-hard on an undirected graph. Only vertex geography on a graph is polynomial time solvable.

# Undirected Vertex Geography

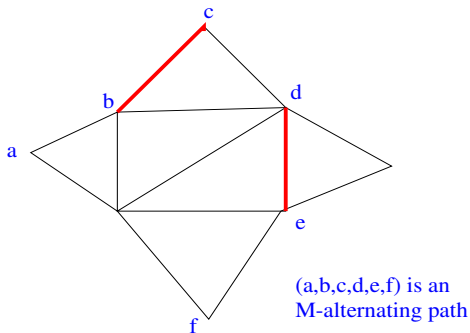
We need some simple results from the theory of matchings on graphs.

A *matching*  $M$  of a graph  $G = (V, E)$  is a set of edges, no two of which are incident to a common vertex.



# Undirected Vertex Geography

$M$ -alternating path



An  $M$ -alternating path joining 2  $M$ -unsaturated vertices is called an  $M$ -augmenting path.

# Undirected Vertex Geography

$M$  is a *maximum* matching of  $G$  if no matching  $M'$  has more edges.

## Theorem

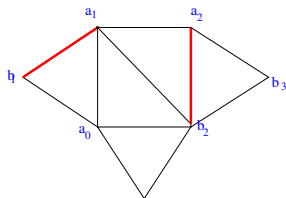
$M$  is a maximum matching iff  $M$  admits no  $M$ -augmenting paths.

**Proof** Suppose  $M$  has an augmenting path

$P = (a_0, b_1, a_1, \dots, a_k, b_{k+1})$  where

$e_i = (a_{i-1}, b_i) \notin M, 1 \leq i \leq k+1$  and

$f_i = (b_i, a_i) \in M, 1 \leq i \leq k.$



Let  $M' = M - \{f_1, f_2, \dots, f_k\} + \{e_1, e_2, \dots, e_{k+1}\}$ .



# Undirected Vertex Geography

- $|M'| = |M| + 1$ .
- $M'$  is a matching

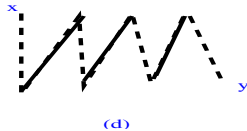
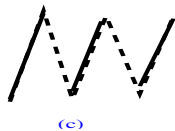
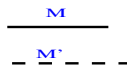
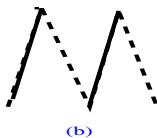
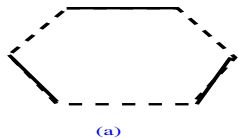
For  $x \in V$  let  $d_M(x)$  denote the degree of  $x$  in matching  $M$ , So  $d_M(x)$  is 0 or 1.

$$d_{M'}(x) = \begin{cases} d_M(x) & x \notin \{a_0, b_1, \dots, b_{k+1}\} \\ d_M(x) & x \in \{b_1, \dots, a_k\} \\ d_M(x) + 1 & x \in \{a_0, b_{k+1}\} \end{cases}$$

So if  $M$  has an augmenting path it is not maximum.

# Undirected Vertex Geography

Suppose  $M$  is not a maximum matching and  $|M'| > |M|$ .  
Consider  $H = G[M \nabla M']$  where  $M \nabla M' = (M \setminus M') \cup (M' \setminus M)$  is the set of edges in *exactly* one of  $M, M'$ .  
Maximum degree of  $H$  is 2 –  $\leq 1$  edge from  $M$  or  $M'$ . So  $H$  is a collection of vertex disjoint alternating paths and cycles.



$x, y$   $M$ -unsaturated

$|M'| > |M|$  implies that there is at least one path of type (d).  
Such a path is  $M$ -augmenting

# Undirected Vertex Geography

## Theorem

$(G, v)$  is an  $N$ -position in UVG iff every maximum matching of  $G$  covers  $v$ .

**Proof** (i) Suppose that  $M$  is a maximum matching of  $G$  which covers  $v$ . Player 1's strategy is now: Move along the  $M$ -edge that contains the current vertex.

If Player 1 were to lose, then there would exist a sequence of edges  $e_1, f_1, \dots, e_k, f_k$  such that  $v \in e_1$ ,  $e_1, e_2, \dots, e_k \in M$ ,  $f_1, f_2, \dots, f_k \notin M$  and  $f_k = (x, y)$  where  $y$  is the current vertex for Player 1 and  $y$  is not covered by  $M$ .

But then if  $A = \{e_1, e_2, \dots, e_k\}$  and  $B = \{f_1, f_2, \dots, f_k\}$  then  $(M \setminus A) \cup B$  is a maximum matching (same size as  $M$ ) which does not cover  $v$ , contradiction.

# Undirected Vertex Geography

(ii) Suppose now that there is some maximum matching  $M$  which does not cover  $v$ . If  $(v, w)$  is Player 1's move, then  $w$

must be covered by  $M$ , else  $M$  is not a maximum matching.

Player 2's strategy is now: Move along the  $M$ -edge that contains the current vertex. If Player 2 were to lose then there exists  $e_1 = (v, w), f_1, \dots, e_k, f_k, e_{k+1} = (x, y)$  where  $y$  is the current vertex for Player 2 and  $y$  is not covered by  $M$ .

But then we have defined an augmenting path from  $v$  to  $y$  and so  $M$  is not a maximum matching, contradiction.  $\square$

# Undirected Vertex Geography

Note that we can determine whether or not  $v$  is covered by all maximum matchings as follows: Find the size  $\sigma$  of the maximum matching  $G$ .

This can be done in  $O(n^3)$  time on an  $n$ -vertex graph. Find the size  $\sigma'$  of a maximum matching in  $G - v$ . Then  $v$  is covered by all maximum matchings of  $G$  iff  $\sigma \neq \sigma'$ .

# Undirected Edge Geography on a bipartite graph

An *even kernel* of  $G$  is a non-empty set  $S \subseteq V$  such that (i)  $S$  is an independent set and (ii)  $v \notin S$  implies that  $\deg_S(v)$  is even, (possibly zero). ( $\deg_S(v)$  is the number of neighbours of  $v$  in  $S$ .)

## Lemma

If  $S$  is an even kernel and  $v \in S$  then  $(G, v)$  is a  $P$ -position in UEG.

**Proof** Any move at a vertex in  $S$  takes the chip outside  $S$  and then Player 2 can immediately put the chip back in  $S$ . After a move from  $x \in S$  to  $y \notin S$ ,  $\deg_S(y)$  will become odd and so there is an edge back to  $S$ . making this move, makes  $\deg_S(y)$  even again. Eventually, there will be no  $S : \bar{S}$  edges and Player 1 will be stuck in  $S$ .  $\square$

# Undirected Edge Geography on a bipartite graph

We now discuss Bipartite UEG i.e. we assume that  $G$  is bipartite,  $G$  has bipartition consisting of a copy of  $[m]$  and a disjoint copy of  $[n]$  and edges set  $E$ . Now consider the  $m \times n$  0-1 matrix  $A$  with  $A(i, j) = 1$  iff  $(i, j) \in E$ .

We can play our game on this matrix: We are either positioned at row  $i$  or we are positioned at column  $j$ . If say, we are positioned at row  $i$ , then we choose a  $j$  such that  $A(i, j) = 1$  and (i) make  $A(i, j) = 0$  and (ii) move the position to column  $j$ . An analogous move is taken when we positioned at column  $j$ .

## Lemma

*Suppose the current position is row  $i$ . This is a P-position iff row  $i$  is in the span of the remaining rows (is the sum (mod 2) of a subset of the other rows) or row  $i$  is a zero row. A similar statement can be made if the position is column  $j$ .*

# Undirected Edge Geography on a bipartite graph

**Proof** If row  $i$  is a zero row then vertex  $i$  is isolated and this is clearly a P-position. Otherwise, assume the position is row 1 and there exists  $I \subseteq [m]$  such that  $1 \in I$  and

$$r_1 = \sum_{i \in I \setminus \{1\}} r_i \pmod{2} \text{ or } \sum_{i \in I} r_i = 0 \pmod{2} \quad (9)$$

where  $r_i$  denotes row  $i$ .

$I$  is an even kernel: If  $x \notin I$  then either (i)  $x$  corresponds to a row and there are no  $x, I$  edges or (ii)  $x$  corresponds to a column and then  $\sum_{i \in I} A(i, x) = 0 \pmod{2}$  from (9) and then  $x$  has an even number of neighbours in  $I$ .



# Undirected Edge Geography on a bipartite graph

Now suppose that (9) does not hold for any  $l$ . We show that there exists a  $l$  such that  $A(1, l) = 1$  and putting  $A(1, l) = 0$  makes column  $l$  dependent on the remaining columns. Then we will be in a P-position, by the first part.

Let  $e_1$  be the  $m$ -vector with a 1 in row 1 and a 0 everywhere else. Let  $A^*$  be obtained by adding  $e_1$  to  $A$  as an  $(n+1)$ th column. Now the row-rank of  $A^*$  is the same as the row-rank of  $A$  (here we are doing all arithmetic modulo 2). Suppose not, then if  $r_i^*$  is the  $i$ th row of  $A^*$  then there exists a set  $J$  such that

$$\sum_{i \in J} r_i = 0(\text{mod } 2) \neq \sum_{i \in J} r_i^*(\text{mod } 2).$$

Now  $1 \notin J$  because  $r_1$  is independent of the remaining rows of  $A$ , but then  $\sum_{i \in J} r_i = 0(\text{mod } 2)$  implies  $\sum_{i \in J} r_i^* = 0(\text{mod } 2)$  since the last column has all zeros, except in row 1.

# Undirected Edge Geography on a bipartite graph

Thus  $\text{rank } A^* = \text{rank } A$  and so there exists  $K \subseteq [n]$  such that

$$e_1 = \sum_{k \in K} c_k \pmod{2} \text{ or } e_1 + \sum_{k \in K} c_k = 0 \pmod{2} \quad (10)$$

where  $c_k$  denotes column  $k$  of  $A$ .

Thus there exists  $\ell \in K$  such that  $A(1, \ell) = 1$ . Now let  $c'_j = c_j$  for  $j \neq \ell$  and  $c'_\ell$  be obtained from  $c_\ell$  by putting  $A(1, \ell) = 0$  i.e.  $c'_\ell = c_\ell + e_1$ . But then (10) implies that  $\sum_{k \in K} c'_k = 0 \pmod{2}$  ( $K = \{k\}$  is a possibility here).. □

# Tic Tac Toe

We consider the following multi-dimensional version of Tic Tac Toe (Noughts and Crosses to the English).

The *board* consists of  $[n]^d$ . A point on the board is therefore a vector  $(x_1, x_2, \dots, x_d)$  where  $1 \leq x_i \leq n$  for  $1 \leq i \leq d$ .

A *line* is a set points  $(x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(d)})$ ,  $j = 1, 2, \dots, n$  where each sequence  $x^{(i)}$  is either (i) of the form  $k, k, \dots, k$  for some  $k \in [n]$  or is (ii)  $1, 2, \dots, n$  or is (iii)  $n, n-1, \dots, 1$ . Finally, we cannot have Case (i) for all  $i$ .

Thus in the (familiar)  $3 \times 3$  case, the top row is defined by  $x^{(1)} = 1, 1, 1$  and  $x^{(2)} = 1, 2, 3$  and the diagonal from the bottom left to the top right is defined by  $x^{(1)} = 3, 2, 1$  and  $x^{(2)} = 1, 2, 3$

## Lemma

The number of winning lines in the  $(n, d)$  game is  $\frac{(n+2)^d - n^d}{2}$ .

**Proof** In the definition of a line there are  $n$  choices for  $k$  in (i) and then (ii), (iii) make it up to  $n + 2$ . There are  $d$  independent choices for each  $i$  making  $(n + 2)^d$ .

Now delete  $n^d$  choices where only Case (i) is used. Then divide by 2 because replacing (ii) by (iii) and vice-versa whenever Case (i) does not hold produces the same set of points (traversing the line in the other direction). □

# Tic Tac Toe

The game is played by 2 players. The Red player (X player) goes first and colours a point red. Then the Blue player (O player) colours a different point blue and so on.

A player wins if there is a line, all of whose points are that player's colour. If neither player wins then the game is a draw. The second player does not have a winning strategy:

## Lemma

*Player 1 can always get at least a draw.*

**Proof** We prove this by considering *strategy stealing*.

Suppose that Player 2 did have a winning strategy. Then Player 1 can make an arbitrary first move  $x_1$ . Player 2 will then move with  $y_1$ . Player 1 will now win playing the winning strategy for Player 2 against a first move of  $y_1$ .

This can be carried out until the strategy calls for move  $x_1$  (if at all). But then Player 1 can make an arbitrary move and continue, since  $x_1$  has already been made. □

The Hales-Jewett Theorem of Ramsey Theory implies that there is a winner in the  $(n, d)$  game, when  $n$  is large enough with respect to  $d$ . The winner is of course Player 1.

# Tic Tac Toe

$$\begin{bmatrix} 11 & 1 & 8 & 1 & 12 \\ 6 & 2 & 2 & 9 & 10 \\ 3 & 7 & * & 9 & 3 \\ 6 & 7 & 4 & 4 & 10 \\ 12 & 5 & 8 & 5 & 11 \end{bmatrix}$$

The above array gives a strategy for Player 2 in the  $5 \times 5$  game ( $d = 2, n = 5$ ).

For each of the 12 lines there is an associated pair of positions. If Player 1 chooses a position with a number  $i$ , then Player 2 responds by choosing the other cell with the number  $i$ .

This ensures that Player 1 cannot take line  $i$ . If Player 1 chooses the \* then Player 2 can choose any cell with an unused number.

# Tic Tac Toe

So, later in the game if Player 1 chooses a cell with  $j$  and Player 2 already has the other  $j$ , then Player 2 can choose an arbitrary cell.

Player 2's strategy is to ensure that after all cells have been chosen, he/she will have chosen one of the numbered cells associated with each line. This prevents Player 1 from taking a whole line. This is called a *pairing* strategy.



# Tic Tac Toe

We now generalise the game to the following: We have a family  $\mathcal{F} = A_1, A_2, \dots, A_N \subseteq A$ . A move consists of one player, taking an uncoloured member of  $A$  and giving it his colour.

A player wins if one of the sets  $A_i$  is completely coloured with his colour.

A pairing strategy is a collection of distinct elements  $X = \{x_1, x_2, \dots, x_{2N-1}, x_{2N}\}$  such that  $x_{2i-1}, x_{2i} \in A_i$  for  $i \geq 1$ .

This is called a *draw forcing pairing*. Player 2 responds to Player 1's choice of  $x_{2i+\delta}$ ,  $\delta = 0, 1$  by choosing  $x_{2i+3-\delta}$ . If Player 1 does not choose from  $X$ , then Player 2 can choose any uncoloured element of  $X$ .

# Tic Tac Toe

In this way, Player 2 avoids defeat, because at the end of the game Player 2 will have coloured at least one of each of the pairs  $x_{2i-1}, x_{2i}$  and so Player 1 cannot have completely coloured  $A_i$  for  $i = 1, 2, \dots, N$ .

## Theorem

If

$$\left| \bigcup_{X \in \mathcal{G}} X \right| \geq 2|\mathcal{G}| \quad \forall \mathcal{G} \subseteq \mathcal{F} \quad (11)$$

then there is a draw forcing pairing.

**Proof** We define a bipartite graph  $\Gamma$ .  $A$  will be one side of the bipartition and  $B = \{b_1, b_2, \dots, b_{2N}\}$ . Here  $b_{2i-1}$  and  $b_{2i}$  both represent  $A_i$  in the sense that if  $a \in A_i$  then there is an edge  $(a, b_{2i-1})$  and an edge  $(a, b_{2i})$ .

A draw forcing pairing corresponds to a complete matching of  $B$  into  $A$  and the condition (11) implies that Hall's condition is satisfied. □

## Corollary

If  $|A_i| \geq n$  for  $i = 1, 2, \dots, n$  and every  $x \in A$  is contained in at most  $n/2$  sets of  $\mathcal{F}$  then there is a draw forcing pairing.

**Proof** The degree of  $a \in A$  is at most  $2(n/2)$  in  $\Gamma$  and the degree of each  $b \in B$  is at least  $n$ . This implies (via Hall's condition) that there is a complete matching of  $B$  into  $A$ .  $\square$

# Tic Tac Toe

Consider Tic tac Toe when  $d = 2$ . If  $n$  is even then every array element is in at most 3 lines (one row, one column and at most one diagonal) and if  $n$  is odd then every array element is in at most 4 lines (one row, one column and at most two diagonals).

Thus there is a draw forcing pairing if  $n \geq 6$ ,  $n$  even and if  $n \geq 9$ ,  $n$  odd. (The cases  $n = 4, 7$  have been settled as draws.  $n = 7$  required the use of a computer to examine all possible strategies.)

# Tic Tac Toe

In general we have

## Lemma

*If  $n \geq 3^d - 1$  and  $n$  is odd or if  $n \geq 2^d - 1$  and  $n$  is even, then there is a draw forcing pairing of  $(n, d)$  Tic tac Toe.*

**Proof** We only have to estimate the number of lines through a fixed point  $\mathbf{c} = (c_1, c_2, \dots, c_d)$ .

If  $n$  is odd then to choose a line  $L$  through  $\mathbf{c}$  we specify, for each index  $i$  whether  $L$  is (i) constant on  $i$ , (ii) increasing on  $i$  or (iii) decreasing on  $i$ .

This gives  $3^d$  choices. Subtract 1 to avoid the all constant case and divide by 2 because each line gets counted twice this way.

# Tic Tac Toe

When  $n$  is even, we observe that once we have chosen in which positions  $L$  is constant,  $L$  is determined.

Suppose  $c_1 = x$  and 1 is not a fixed position. Then every other non-fixed position is  $x$  or  $n - x + 1$ . Assuming w.l.o.g. that  $x \leq n/2$  we see that  $x < n - x + 1$  and the positions with  $x$  increase together at the same time as the positions with  $n - x + 1$  decrease together.

Thus the number of lines through  $\mathbf{c}$  in this case is bounded by  $\sum_{i=0}^{d-1} \binom{d}{i} = 2^d - 1$ . □

# Quasi-probabilistic method

We now prove a theorem of Erdős and Selfridge.

## Theorem

If  $|A_i| \geq n$  for  $i \in [N]$  and  $N < 2^{n-1}$ , then Player 2 can get a draw in the game defined by  $\mathcal{F}$ .

**Proof** At any point in the game, let  $C_j$  denote the set of elements in  $A$  which have been coloured with Player  $j$ 's colour,  $j = 1, 2$  and  $U = A \setminus C_1 \cup C_2$ . Let

$$\Phi = \sum_{i: A_i \cap C_2 = \emptyset} 2^{-|A_i \cap U|}.$$

Suppose that the players choices are  $x_1, y_1, x_2, y_2, \dots$ . Then we observe that immediately after Player 1's first move,  $\Phi < N2^{-(n-1)} < 1$ .



# Quasi-probabilistic method

We will show that Player 2 can keep  $\Phi < 1$  through out. Then at the end, when  $U = \emptyset$ ,  $\Phi = \sum_{i:A_i \cap C_2 = \emptyset} 1 < 1$  implies that  $A_i \cap C_2 \neq \emptyset$  for all  $i \in [N]$ .

So, now let  $\Phi_j$  be the value of  $\Phi$  after the choice of  $x_1, y_1, \dots, x_j$ . then if  $U, C_1, C_2$  are defined at precisely this time,

$$\begin{aligned}\Phi_{j+1} - \Phi_j &= - \sum_{\substack{i:A_i \cap C_2 = \emptyset \\ y_j \in A_i}} 2^{-|A_i \cap U|} + \sum_{\substack{i:A_i \cap C_2 = \emptyset \\ y_j \notin A_i, x_{j+1} \in A_i}} 2^{-|A_i \cap U|} \\ &\leq - \sum_{\substack{i:A_i \cap C_2 = \emptyset \\ y_j \in A_i}} 2^{-|A_i \cap U|} + \sum_{\substack{i:A_i \cap C_2 = \emptyset \\ x_{j+1} \in A_i}} 2^{-|A_i \cap U|}\end{aligned}$$

# Quasi-probabilistic method

We deduce that  $\Phi_{j+1} - \Phi_j \leq 0$  if Player 2 chooses  $y_j$  to maximise  $\sum_{\substack{i: A_i \cap C_2 = \emptyset \\ y \in A_i}} 2^{-|A_i \cap U|}$  over  $y$ .

In this way, Player 2 keeps  $\Phi < 1$  and obtains a draw. □

In the case of  $(n, d)$  Tic Tac Toe, we see that Player 2 can force a draw if

$$\frac{(n+2)^d - n^d}{2} < 2^{n-1}$$

which is implied, for  $n$  large, by

$$n \geq (1 + \epsilon)d \log_2 d$$

where  $\epsilon > 0$  is a small positive constant.