that actual games hardly ever last to the boundary of the chequered exercise book page they are customarily played on, not even close. Still for the "theory" of the game, this extension of the board size makes a huge difference: for one game we know the solution, for the other we do not. The 15-by-15 `FP` winning strategy of course implies that Maker wins in the Maker-Breaker version of the game. This in turn implies that Maker also has a winning strategy on the infinite board. Curiously, it is still **not** known whether `FP` can win the strong 5-in-a-row game on the infinite board. Why cannot `FP` just apply Allis' 15-by-15 `FP`-strategy out of the box to occupy his winning set on the infinite board? Partly because of the dual task of simultaneous offense/defense `FP` has to perform in a strong game. Allis' strategy will certainly create a 5-in-a-row eventually, while preventing `SP` from creating his own on the 15-by-15 board, but along the way it might also lead to a 4-in-a-row of `SP` at the boundary of the 15-by-15 board. This would not concern `FP` in the finite game, but in the infinite game he would be forced to play outside the 15-by-15 board and abandon his winning strategy. This is yet another manifestation of the extra set paradox of the first chapter, or rather of the difficulties it can cause.
It is widely believed that the 5-in-row game on the infinite board is `FP`'s win, while 6-in-a-row is a draw. However, it is only known that 4-in-a-row is `FP`'s win (a trivial exercise) and that 8-in-a-row is a draw. In Section 2.4.2 we will see a proof that 40-in-a-row is a draw.

2.2 Coloring hypergraphs

Let $\binom{X}{k} := \{K \subseteq X : |K| = k\}$ the set of all k-element subsets of X.
A hypergraph $(X, \mathcal{F})$ is called *k-uniform* if $\mathcal{F} \subseteq \binom{X}{k}$ consists only of k-element subsets. Sometimes we identify the hypergraph with its edge set $\mathcal{F}$.

A function $f : X \to \{\texttt{red}, \texttt{blue}\}$ is called a *proper 2-coloring* of the hypergraph $(X, \mathcal{F})$ if every member of $\mathcal{F}$ has both a `red` and a `blue` colored vertex (that is, no edge is *monochromatic*). A hypergraph $(X, \mathcal{F})$ is called *2-colorable* if it has a proper 2-coloring.

For a proper 2-coloring to exist we obviously need at least two vertices in each edge, so we assume $k \geq 2$. For example a k-uniform hypergraph with exactly two distinct edges always has a proper 2-coloring: in each of the two edges take a vertex which is not part of the other edge and color it `red`, and color the rest of the vertices `blue`. It is a famous open problem to determine for each k the smallest number $m(k)$ of edges in a **non**-2-colorable k-uniform hypergraph. The triangle graph shows that $m(2) = 3$.

---

game the players' goal is still to be the first to occupy a winning set fully, and a draw means that none of the players wins, hence the game goes on infinitely long. Strategy stealing still applies, so `SP` can only hope for a draw. In the Maker-Breaker game, Breaker wins if Maker does not occupy a winning set. This is not any more equivalent to Breaker putting his mark in every winning set: the point is that Breaker can force the game to last infinitely long without Maker winning.

The following result of Erdős [27], one of the first applications of the probabilistic method, provides an exponential lower bound.

Claim 2.2.1. *If* $|\mathcal{F}| < 2^{k-1}$*, then* $\mathcal{F}$ *is* 2*-colorable. In particular,* $m(k) \geq 2^{k-1}$*.*

Proof. Take a random 2-coloring $f : V(\mathcal{F}) \to \{\texttt{red}, \texttt{blue}\}$. That is: color all vertices $x \in V(\mathcal{F})$ independently, uniformly at random such that

$$\Pr[f(x) = \texttt{red}] = \frac{1}{2} = \Pr[f(x) = \texttt{blue}].$$

For each $A \in \mathcal{F}$, let Y_A be the *characteristic random variable* of the event that A is monochromatic. That is $Y_A = 1$ if A is monochromatic, otherwise $Y_A = 0$. Now,

$$\mathbb{E}[\#\text{of monochromatic edges of } \mathcal{F}] = \mathbb{E}\left[\sum_{A \in \mathcal{F}} Y_A\right] = \sum_{A \in \mathcal{F}} \mathbb{E}\left[Y_A\right] = \frac{|\mathcal{F}|}{2^{k-1}} < 1.$$

The random variable $\sum_{A \in \mathcal{F}} Y_A$ takes only non-negative integer values, so its average can only be strictly less than 1 if it also takes the value 0 at least once. Hence for sure, not just with some probability, but with 100% certainty, **there exists** a 2-coloring of $\mathcal{F}$ without monochromatic edges. □

Remark 2.2.2. The classical term for a hypergraph $\mathcal{F}$ being 2-colorable is that $\mathcal{F}$ has *property B*.

Remark 2.2.3. The best known lower bound for $m(k)$, due to Radhakrishnan and Srinivasan [76], is of the order $2^k\sqrt{k/\log k}$. Their proof also starts with a random coloring of the vertices, but continues with a refined randomized recoloring procedure, which fixes the errors (the monochromatic sets). The best known upper bound, obtained by considering a random k-uniform hypergraph on roughly k^2 vertices, is of the order $2^k k^2$. It was proved by Paul Erdős [28] around the same time "Oh Pretty Woman" topped the US charts: not exactly yesterday. Improving these bounds remain intriguing open problems.

The probabilistic method is a great tool to prove the *existence* of special objects, including, as the case may be, a coloring of the vertices of a hypergraph with `red` and `blue` so that in each hyperedge both `red` and `blue` occur. In our computer-oriented world, however, a proof of the existence of something is of little value: We do not just want to *know* that this book exists and that some people actually own a copy, but we want to hold one in our hands. With such worldly possessions, this is usually easy to accomplish with enough money at hand. However, in the case of a proper coloring of a hypergraph, the existence of which is proved by the above claim, we need something else.

When do we hold a 2-coloring of a hypergraph in our hand? We of course want that, given the hypergraph in some form (say, by its incidence matrix) we can *construct* a proper coloring, potentially with the help of a computer. So we need an *algorithm* for this task. This is easy, you might say, once we know the coloring exists: the number of vertices is finite, the hypergraph is finite, so we could look at all possible 2-colorings of the vertex set and check for each whether it is proper. This solution, however, is not really satisfying, as the number of all colorings is $2^{|X|}$ and no matter how super a computer you own now (or will own, ever) it will never even get close to finishing the check within the lifetime of our universe, on even relatively small problems where the vertex set X is, say, of size 100.

In computer science one usually accepts algorithms as theoretically "fast" when their running time is polynomial in the size of the input. Is there a polynomial time algorithm which *finds* a proper 2-coloring of any given input hypergraph from Claim 2.2.1?

Based on the proof of Claim 2.2.1, we can suggest a natural randomized algorithm: Color the vertices independently, uniformly at random and then check whether the obtained coloring is proper. If the answer is YES, then output this coloring and terminate, otherwise repeat the procedure with a new random coloring. Generating a random assignment involves $|X|$ calls to a uniformly random source and the checking involves going through the colors of the k vertices of each of the $|\mathcal{F}|$ hyperedges. Altogether this represents $|X| + k|\mathcal{F}| = O(|X| \cdot |\mathcal{F}|)$ steps, which is just quadratic in the input data.

Assuming $|\mathcal{F}| < 2^{k-2}$, only a factor 2 stronger assumption than the one in Claim 2.2.1, we find that the probability of failure of the random coloring producing a proper coloring is at most $\frac{1}{2}$, by Markov's Inequality. After just one hundred iterations, which shows up in the running time just as a multiplicative constant factor 100, the probability of failure to produce a proper coloring is at most $\frac{1}{2^{100}}$, orders of magnitude smaller than the probability of hardware failure of your supercomputer resulting in an incorrect outcome.

Although, for all practical purposes, such an algorithm is satisfactory, still there is the dependence on a perfectly uniform random source and the loss of the constant factor 2 compared to the existence result of the claim. The question remains whether there is an efficient and possibly deterministic way to find a proper coloring whose existence is promised by Claim 2.2.1.

The answer to this question is also YES and it was given in the context of positional games by Erdős and Selfridge.

2.3 The Erdős-Selfridge Criterion

The following proposition establishes the connection between 2-colorings and Maker-Breaker games. The argument is analogous to the one we saw in the proof of Theorem 1.3.2 of Chapter 1.

Proposition 2.3.1. *$\mathcal{F}$ is a Breaker's win $\Rightarrow$ $\mathcal{F}$ is 2-colorable.*

Proof. Let us sit two players, FP and SP, to play on the board $V(\mathcal{F})$ and give both of them the winning strategy S of Breaker for the game $\mathcal{F}$ (which exists by assumption). More precisely, give S to SP, as Breaker plays second in a Maker-Breaker game, and give FP the winning strategy of Breaker as a first player, which exists by Proposition 2.1.6. Make both players play according to this strategy, such that FP colors his board elements with red and SP colors his with blue. Since FP plays according to a strategy which is a winning one for Breaker, he will put a red mark in every winning set by the end of the game. SP also plays according to Breaker's winning strategy and hence by the end of the game a blue vertex as well will be in every winning set. Looking at the board at the end of the game, we see that every winning set contains both red and blue vertices: the coloring created by the two players during their play is proper. □

By this proposition, the following fundamental result of Erdős and Selfridge is a strengthening of Claim 2.2.1. From its proof it will be obvious how to devise a very efficient deterministic algorithm which finds a proper 2-coloring of the underlying hypergraph.

Theorem 2.3.2. *Let $\mathcal{F}$ be a k-uniform hypergraph. Then*

$$|\mathcal{F}| < 2^{k-1} \;\Rightarrow\; \mathcal{F} \text{ is a Breaker's win.}$$

This theorem is a corollary of the following more general one, dealing not only with uniform hypergraphs.

Theorem 2.3.3 (Erdős-Selfridge Criterion, [31])**.** *Let $\mathcal{F}$ be a hypergraph. Then*

$$\sum_{A\in\mathcal{F}} 2^{-|A|} < \frac{1}{2} \quad\Rightarrow\quad \mathcal{F} \text{ is a Breaker's win.}$$

Proof. Imagine yourself in the middle of a play when Breaker must decide which unoccupied (i.e., uncolored) element of the board to take (i.e., color with blue). Each winning set which does not yet contain Breaker's blue mark represents a potential danger for Breaker. The more elements Maker has already colored red from this winning set, the larger danger it represents for Breaker. Having no better idea, our motivation for the *quantitative* danger of a hyperedge comes from a probabilistic view: instead of clever players coloring the elements, we imagine that each remaining element will be colored with red or blue uniformly at random. We set the current danger value of A to be the probability that A is then fully occupied by Maker (i.e., monochromatic red). Hence the danger value of a winning set A not yet touched by Breaker is defined to be $2^{-(\#\text{ of unoccupied elements in } A)}$, while the danger of those already having a blue vertex is 0. We define the *danger* of a hypergraph $\mathcal{H}$ as the sum of the dangers of its edges, that is,

$$\texttt{danger}(\mathcal{H}) = \sum_{\substack{A\in\mathcal{H}\\ A\cap B=\emptyset}} 2^{-|A\setminus M|},$$

where $M \subseteq X$ and $B \subseteq X$ are the sets of vertices occupied at the moment by Maker and Breaker, respectively. Note that this is exactly the expected number of monochromatic `red` edges of $\mathcal{H}$ after a random 2-coloring of the vertices in $X \setminus (M \cup B)$. By our condition, at the beginning of the game

$$\texttt{danger}(\mathcal{F}) = \sum_{A \in \mathcal{F}} 2^{-|A|} < \frac{1}{2}.$$

Let $M_i = \{m_1, \ldots, m_i\}$ be the set of vertices of Maker after round i and let $B_{i-1} = \{b_1, \ldots, b_{i-1}\}$ be the set of vertices of Breaker after round $i-1$. The actual multihypergraph of interest immediately after Maker's ith move has

- board $X_i = X \setminus (B_{i-1} \cup M_i)$ and
- family of winning sets $\mathcal{F}_i = \{A \setminus M_i : A \in \mathcal{F}, A \cap B_{i-1} = \emptyset\}$.

These are the still available vertices of the board and the leftover winning sets which do not yet contain a mark of Breaker. We would like to emphasize that $\mathcal{F}_i$ is a multiset: For each member $A \in \mathcal{F}$ with the property that $A \cap B_{i-1} = \emptyset$ we create a member $A \setminus M_i \in \mathcal{F}_i$ (even if we create the same set more than once).

With his first move, Maker increases the `danger` of each edge containing m_1 by a factor of 2, while the rest of the edges keep their old `danger` value. Hence

$$\texttt{danger}(\mathcal{F}_1) < 1.$$

Breaker will try to adhere to the simple goal of keeping the danger below 1. His strategy is to be as greedy as possible. In each round i, he will occupy a vertex b_i whose occupation decreases the `danger` of the hypergraph $\mathcal{F}_i$ the most. Formally, this is a vertex $b_i \in X_i$ which maximizes $\sum_{z \in A \in \mathcal{F}_i} 2^{-|A|}$ over all vertices $z \in X_i$, because this is the contribution to the total hypergraph danger of exactly those edges whose danger value will be zeroed after Breaker takes his vertex. (In case there are several eligible vertices, Breaker picks one arbitrarily.)

After Maker also takes his vertex m_{i+1}, the danger of the hypergraph increases, because all those edges which contain Maker's choice double their danger. Then the overall change in the danger of the hypergraph is the following:

$$\begin{aligned}
\texttt{danger}(\mathcal{F}_{i+1}) &= \texttt{danger}(\mathcal{F}_i) - \sum_{\substack{A \in \mathcal{F}_i \\ b_i \in A}} 2^{-|A|} + \sum_{\substack{A \in \mathcal{F}_i \\ m_{i+1} \in A}} 2^{-|A|} - \sum_{\substack{A \in \mathcal{F}_i \\ b_i, m_{i+1} \in A}} 2^{-|A|} \\
&\le \texttt{danger}(\mathcal{F}_i) - \sum_{\substack{A \in \mathcal{F}_i \\ b_i, m_{i+1} \in A}} 2^{-|A|} \\
&\le \texttt{danger}(\mathcal{F}_i).
\end{aligned}$$

Notice that in the first equality we added the danger values of all those edges which contain m_{i+1} in order to double their current danger values. This we should not have done, however, for those edges which also contain b_i, since they already have danger 0 after Breaker's move. We correct this error by subtracting the danger of those edges which contain both m_{i+1} and b_i. The first inequality then follows because of how b_i was chosen: m_{i+1} was still available for Breaker to choose in round i, but Breaker still chose b_i, so the sum of the dangers of the edges containing b_i must have been at least as large as the sum of the dangers of those edges containing m_{i+1}.

So if Breaker follows his strategy, for every i we have

$$\texttt{danger}(\mathcal{F}_i) \leq \texttt{danger}(\mathcal{F}_1) < 1. \tag{2.1}$$

If Maker still won the game, say in round i, then we would have $\emptyset \in \mathcal{F}_i$ by the definition of $\mathcal{F}_i$. Now this alone would have contributed $2^{-|\emptyset|} = 1$ to the `danger` of $\mathcal{F}_i$, a contradiction to (2.1).

So playing according to this strategy, Breaker must have won the game. □

Remark 2.3.4. It is easy to see that the greedy strategy of the above proof also gives an efficient deterministic algorithm for Breaker to determine his next move. In each round he would have to calculate $\sum_{z \in A \in \mathcal{F}_i} 2^{-|A|}$ for each $z \in X_i$ and choose the largest one. This is the checking of at most $|X|$ sums of at most $|\mathcal{F}|$ terms each, a calculation of order $|X| \cdot |\mathcal{F}|$ steps, polynomial in the input size.

Remark 2.3.5. The general method of the proof (taking expectation conditioned on the current situation) is of fundamental importance in theoretical computer science. It is the first instance of the *method of conditional expectations*, the very first technique to efficiently *derandomize* randomized algorithms, which is applicable in many, much more general scenarios (see, e.g., [3, 68]).

Remark 2.3.6. The following construction shows that the Erdős-Selfridge Criterion (Theorem 2.3.2) is best possible for every positive integer k. It describes a k-uniform hypergraph with exactly 2^{k-1} edges which is a Maker's win. Let $X = \{r\} \cup \{L_1, \dots, L_{k-1}\} \cup \{R_1, \dots, R_{k-1}\}$ and let $\mathcal{F}$ contain all subsets of size k which

(1) contain the vertex r, and

(2) contain exactly one element of each pair $\{L_i, R_i\}$.

This is a k-uniform hypergraph with exactly 2^{k-1} edges. Maker can win as follows. To start, he takes the vertex r and then in the next $k-1$ moves he acts according to a pairing strategy. That is, he always takes the sibling of what Breaker took previously: if Breaker took L_i, Maker takes R_i, if Breaker took R_i, Maker takes L_i. At the end then Maker owns one of each pair $\{L_i, R_i\}$ as well as the element r. Since each set of such type is in the family $\mathcal{F}$, Maker won the game.

An alternative interpretation of this construction uses a full binary tree of depth $k-1$. The vertex set of the tree is the board and the winning sets are the vertex sets of the 2^{k-1} root-to-leaf paths. Maker starts by taking the root vertex r and then builds a path by always taking a child of his previously taken vertex v, which does *not* contain any vertices of Breaker in its subtree. By induction at least one of the children of v is such, and hence Maker occupies a root-to-leaf path in k moves. The choices L_i, R_i of the previous construction translate in round i for Maker to take the left or right child of v, respectively. Breaker taking a vertex in round i which is *not* in the right subtree of v can be interpreted as if he took L_i, while otherwise he took R_i.

Remark 2.3.7. The converse of the Erdős-Selfridge Theorem is trivially not true: if $\mathcal{F}$ is a k-uniform hypergraph and $|\mathcal{F}| \geq 2^{k-1}$, then it is not necessarily the case that $\mathcal{F}$ is a Maker's win. (Just consider 2^{k-1} disjoint hyperedges.) Some sort of a weaker converse, involving a couple of extra factors, is formulated in the next theorem.

Theorem 2.3.8 (Maker's Win Criterion; Beck)**.** *If $\mathcal{F}$ is a k-uniform hypergraph, then*

$$|\mathcal{F}| > 2^{k-3} \cdot \Delta_2(\mathcal{F}) \cdot |X| \;\Rightarrow\; \mathcal{F} \text{ is a Maker's win,}$$

where $\Delta_2(\mathcal{F}) = \max\{\deg(x,y) : x, y \in X, x \neq y\}$ *and* $\deg(x,y) = |\{A \in \mathcal{F} : x, y \in A\}|$.

Proof. Exercise. □

2.4 Applications of the Erdős-Selfridge Criterion

2.4.1 Clique Game

In the *clique building game* $\mathcal{K}_{K_q}(n)$ the board is the edge set $E(K_n) = \binom{[n]}{2}$ of the complete graph on the vertex set $[n] := \{1, 2, \ldots, n\}$ and the family of winning sets is $\left\{\binom{Q}{2} : Q \subseteq [n], |Q| = q\right\}$, containing the edge sets of q-cliques. Let $q(n)$ be the largest integer q such that Maker wins $\mathcal{K}_{K_q}(n)$.

Theorem 2.4.1.

$$\frac{1}{2}\log_2 n \leq q(n) \leq 2\log_2 n.$$

Proof. If $n \geq R(q,q)$, then, as discussed in the first chapter, not only Maker, but also `FP` of the strong game has a winning strategy. Indeed, draw is not an option because the board will contain a monochromatic clique of order q, and Strategy Stealing shows that `SP` does not have a winning strategy. In particular, Maker can build a clique of size $\frac{1}{2}\log_2 n$, since $R(q,q) < 4^q$.

For the other direction, we use the Erdős-Selfridge Criterion to show that Breaker has a winning strategy. The hypergraph of the clique game $\mathcal{K}_{K_q}(n)$ is $\binom{q}{2}$-uniform and has $\binom{n}{q}$ edges. Substituting into the Erdős-Selfridge Criterion, we get that Breaker wins if $\binom{n}{q} < 2^{\binom{q}{2}-1}$, that is, for example, if

$$\frac{ne}{q} \leq 2^{\frac{q-1}{2}-\frac{1}{q}}.$$