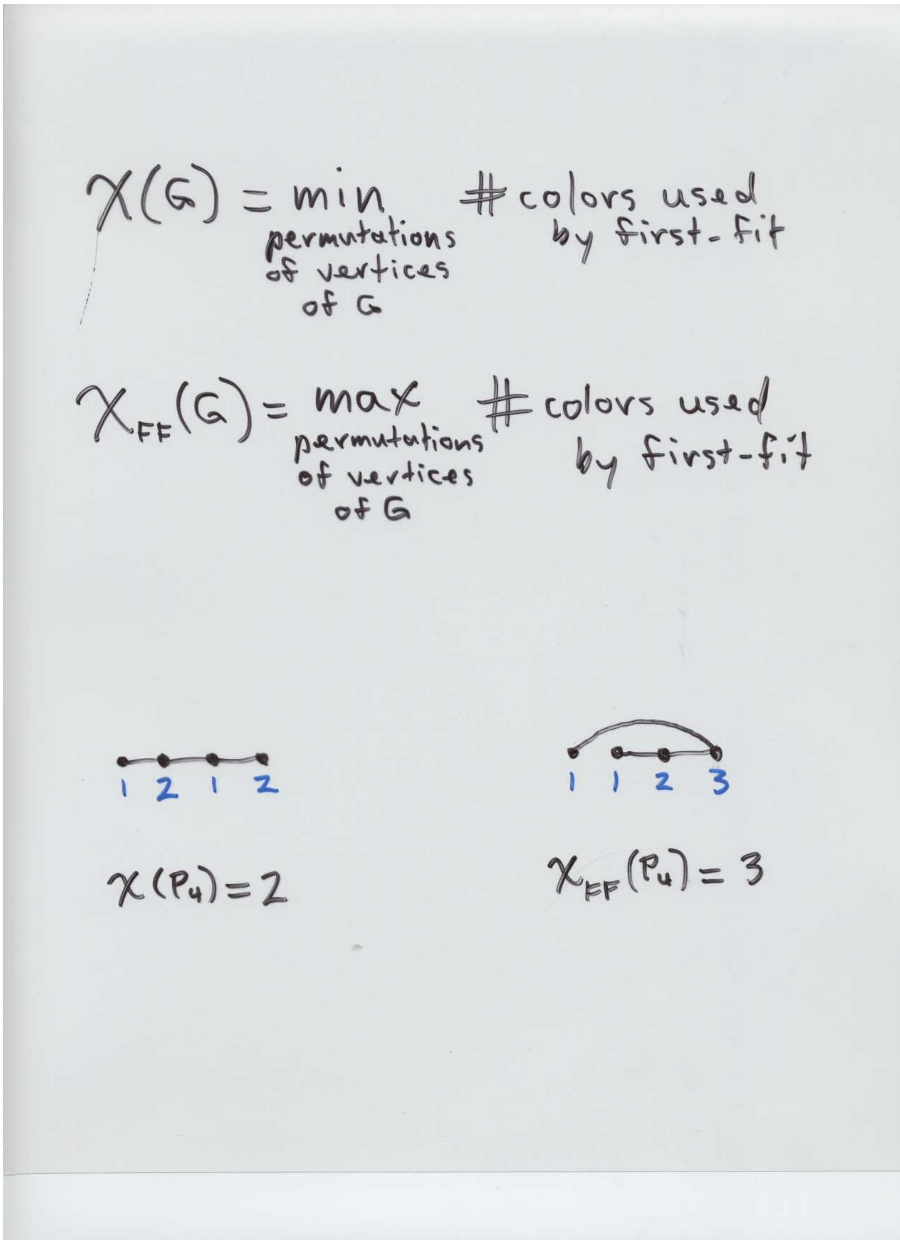
FIRST-FIT COLORING OF INTERVAL GRAPHS

David Smith
with
H. A. Kierstead } Arizona State U.
W. T. Trotter, Georgia Tech

June 2010

These slides are intended for those familiar with graphs and graph coloring.

Slide 1

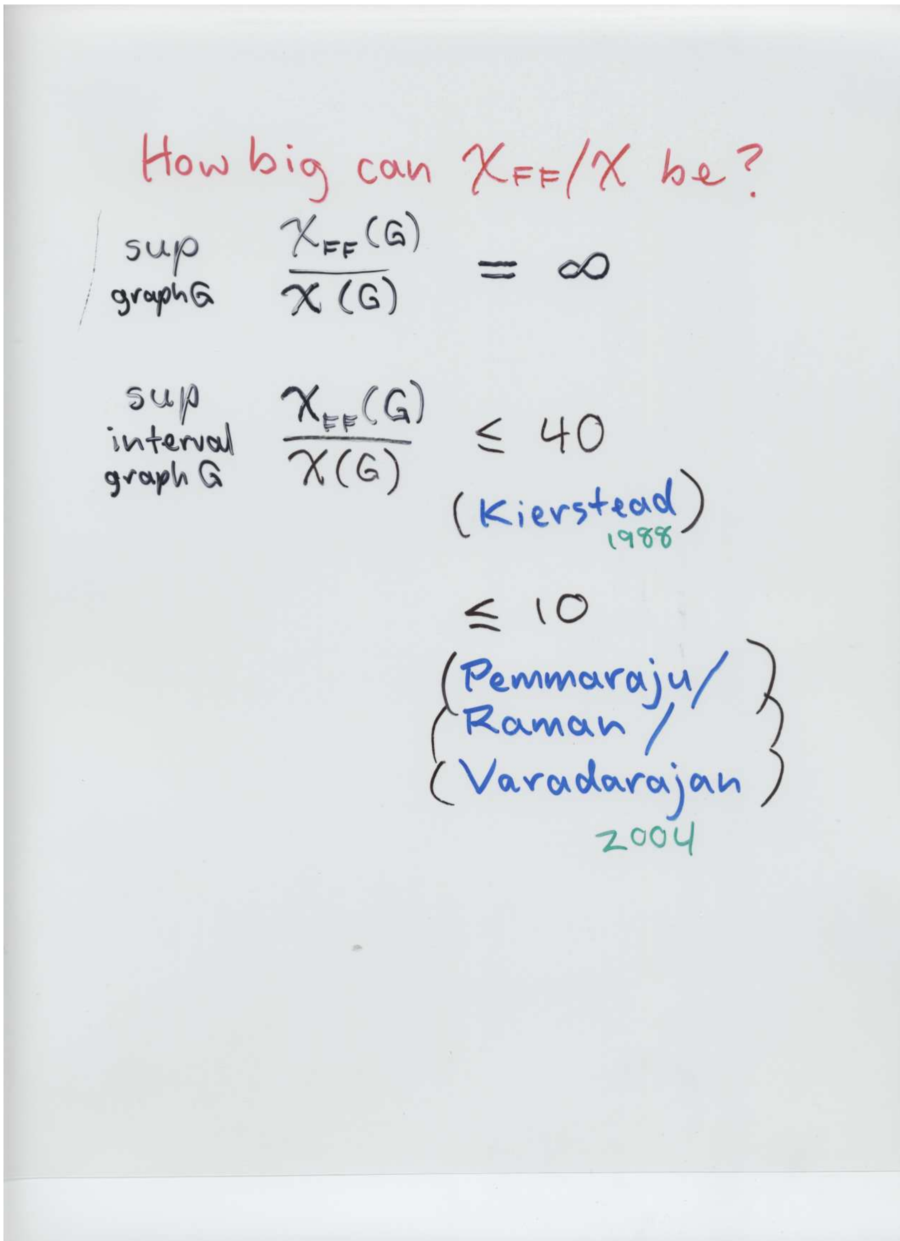


Ordinarily a graph G is considered to be (properly) colored when each vertex has one color, and the colors of neighboring vertices are distinct. The smallest number of colors in such an assignment is $\chi(G)$. The usual goal when coloring is to use few colors.

A simple algorithm that conserves colors is *first-fit*: vertices are arranged in some order, and the algorithm assigns to each vertex the smallest positive integer not already assigned to some neighbor. It is easy to prove that first-fit uses $\chi(G)$ colors when the vertices of G are ordered favorably. We are concerned with the number $\chi_{FF}(G)$ of colors used by first-fit when the vertices of G are presented in a worst order.

For example, the first-fit algorithm uses 2 colors on a path of 4 vertices in the best case. However, 3 colors are required if the ends of the path are presented first.

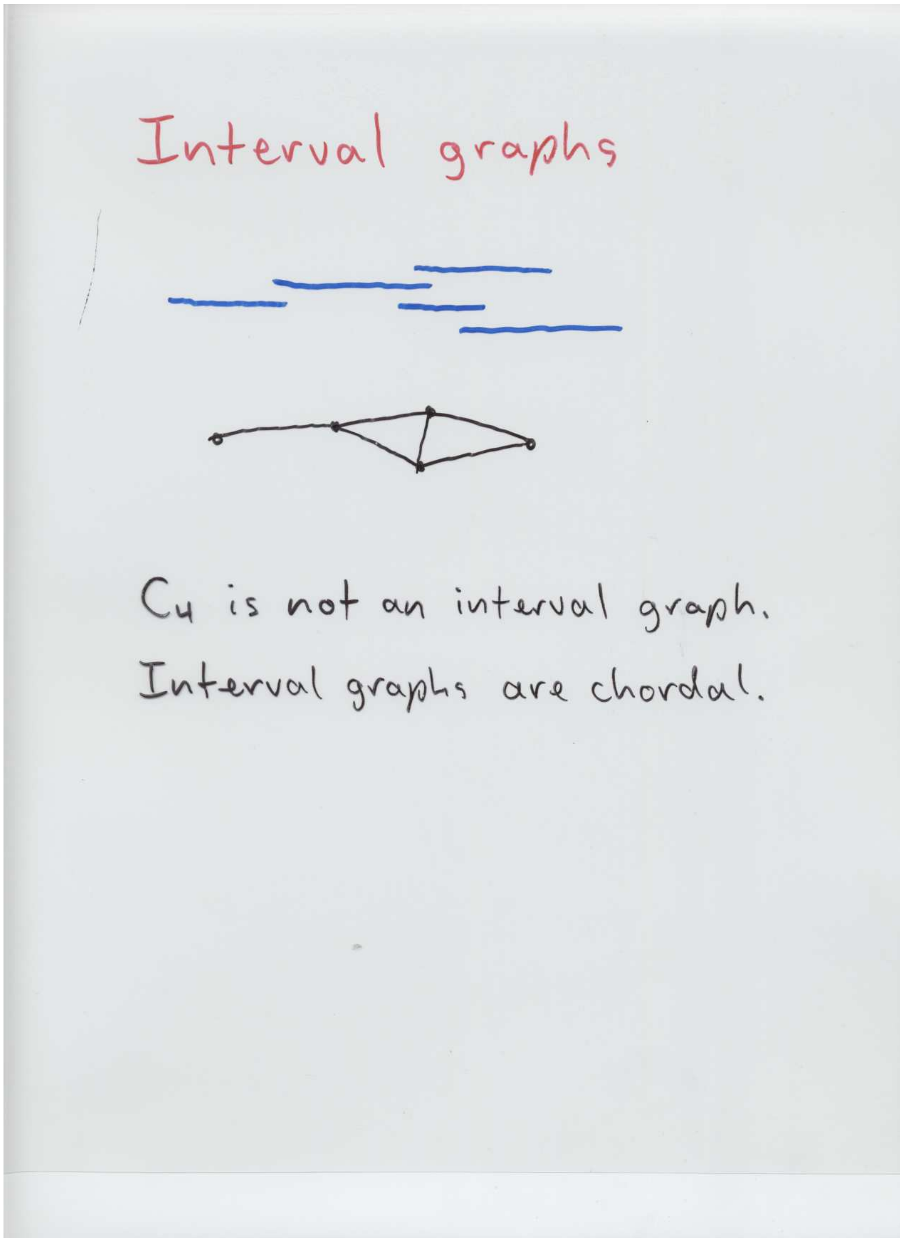
Slide 2



First-fit may use arbitrarily many colors on bipartite graphs, and even on trees. But it uses no more than 40 times the optimal number of colors on interval graphs. This result of Kierstead is the first of its kind. (The history of the problem of first-fit coloring of interval graphs begins perhaps in the late 1960s, but some of it is omitted here for the sake of brevity.) He and his student Qin improved this to 26 or so.

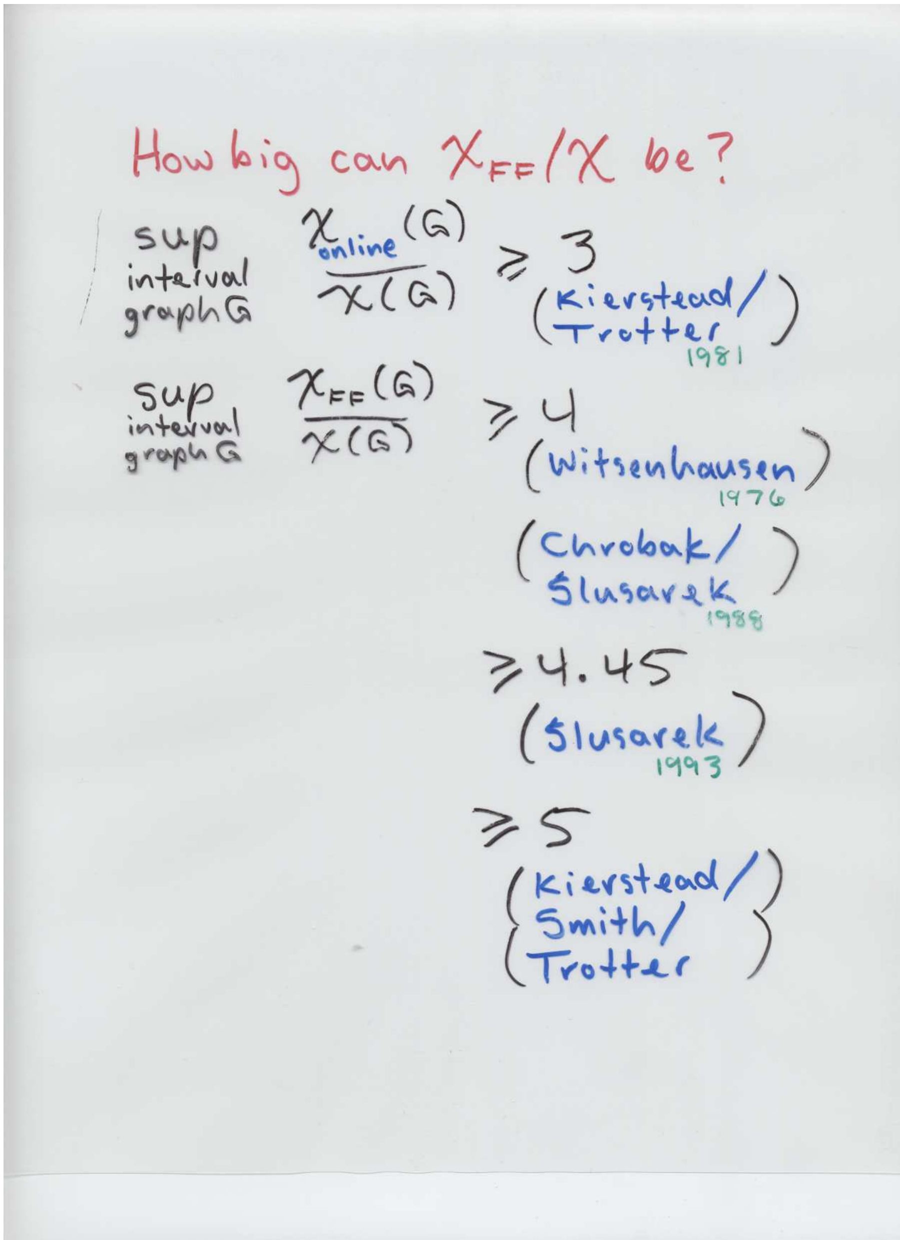
A new approach involving a "column construction procedure" was introduced by Pemmaraju, Raman, and Varadarajan for an upper bound of 10. They nearly obtained 8, and this was completed by Brightwell, Kierstead, and Trotter soon after; and by Narayanaswamy and Subhash Babu, whose proof yields a certain improvement. I learned from Kevin Milans that a similar advance was made recently, but I don't recall whose it is. In any case, the best upper bound remains 8 by the measure $\sup_{\text{int. graph } G} \dfrac{\chi_{FF}(G)}{\chi(G)}$.

Slide 3



Interval graphs, by the way, are defined as having for each vertex a real interval; a pair of vertices is an edge if and only if the intervals meet. This is a proper subclass of the class of graphs. For example, a cycle on 4 vertices is not an interval graph. In fact, interval graphs are chordal (also known as triangulated), so are perfect. When G is perfect, $\chi(G)$ equals the number of vertices in a largest complete subgraph of G.

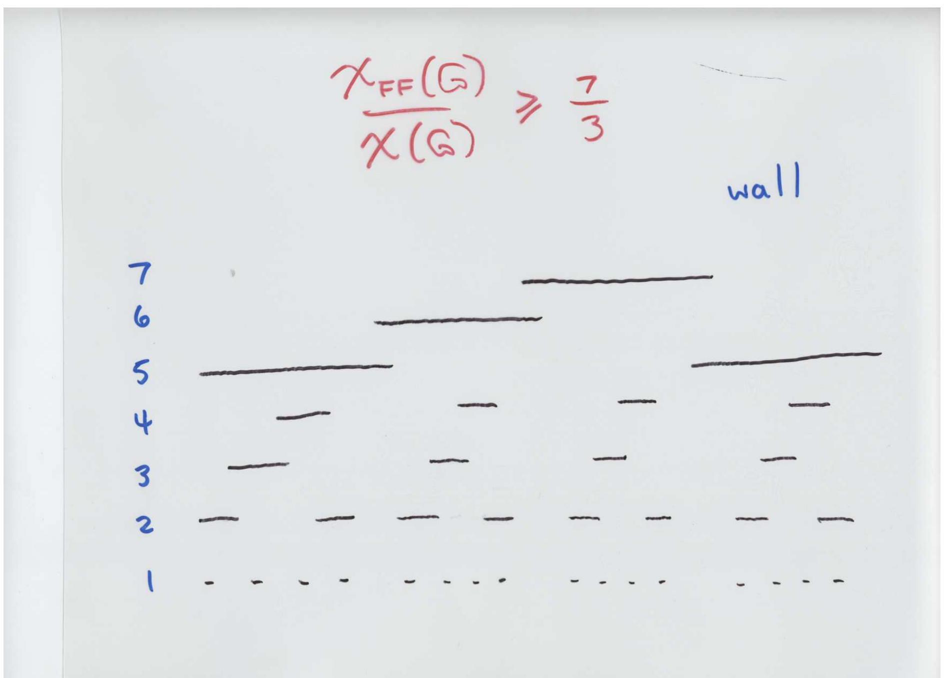
Slide 4



As in the case of a path on 4 vertices, a bad order of the vertices of graph G may cause first-fit to use more than $\chi(G)$ colors. An algorithm that assigns each vertex a color without regard to neighbors not yet presented (as first-fit does) is *online.* Kierstead and Trotter showed that every online algorithm can be made to use asymptotically 3 times the optimal number of colors by construction of a sequence of graphs with suitable vertex orders.

Some years before, Witsenhausen had found for the first-fit algorithm a lower bound of 4. This was rediscovered by Chrobak and Ślusarek. Ślusarek improved the bound to 4.45. A lower bound of 5 was discovered late in 2009. Here are explained the ideas of the proof.

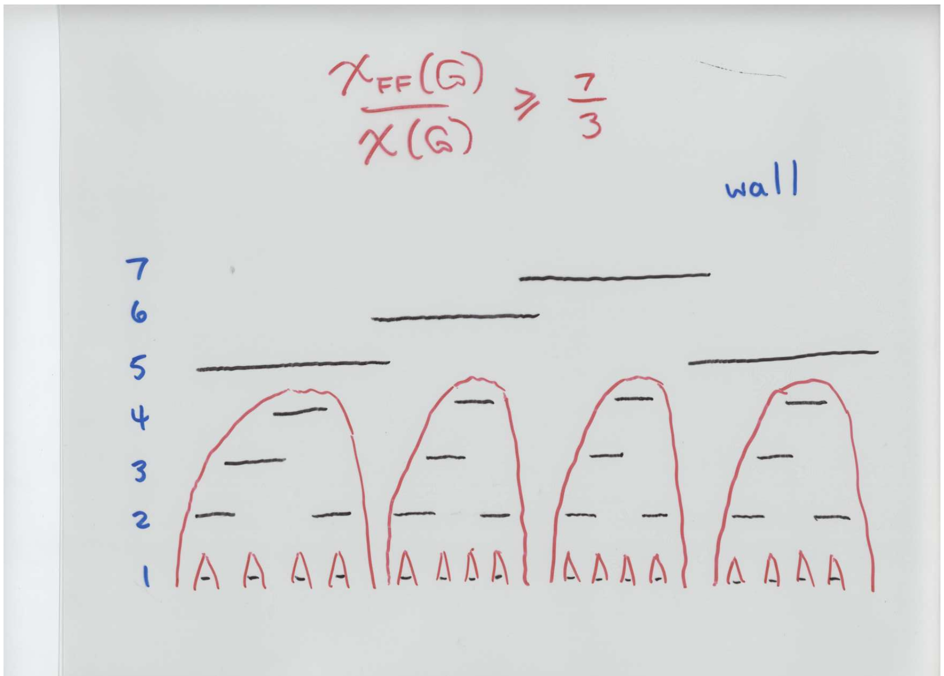
Slide 5



This slide, for example, proves a lower bound of 7/3. First observe that this collection of intervals, seen as an interval graph G, has no complete subgraph larger than a triangle. That is, $\chi(G) \leq 3$.

On the other hand, the intervals are arranged in levels to show that $\chi_{FF}(G) \geq 7$. When vertices are presented in levels from bottom to top, the vertices of level 1 all get color 1; vertices of level 2 all get color 2; and so on. This follows from the facts that each level is an independent set; each vertex of level k has when it is colored a neighbor of each color $1, \ldots, k-1$ (cf. Grundy coloring); and level 7 is occupied by a vertex.

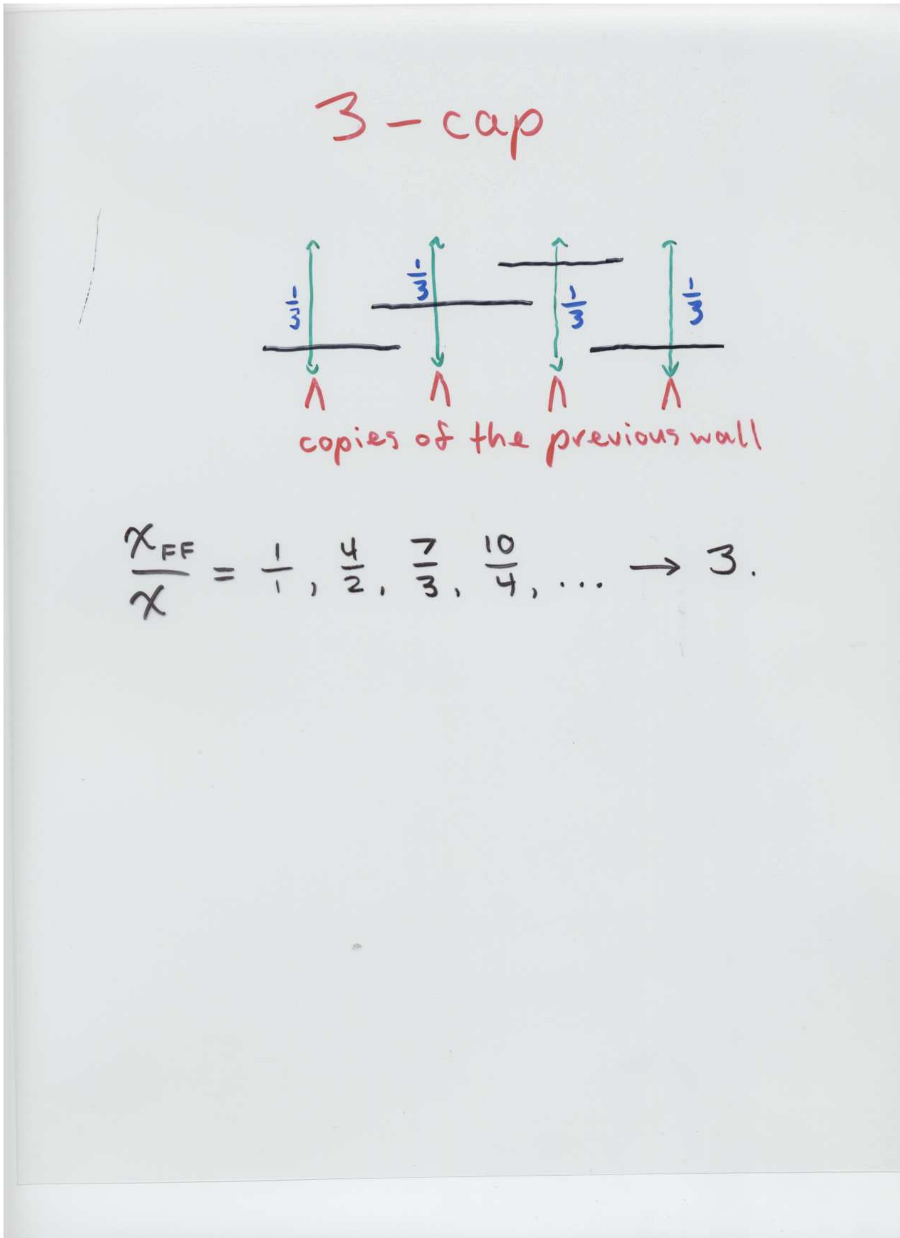
Slide 6



Such a construction, consisting of an interval graph and an assignment of levels, can be imagined as a wall, where each brick is supported not according to some physical condition, but to the Grundy condition.

There is a certain self-similarity in the given example. At the top are 4 intervals arranged in a pattern, beneath each of which are another 4. (Now look at the cap from the bottom up.) At the bottom are 16 little walls each of height 1 and largest clique size 1. In the middle are 4 walls each of height 4 and largest clique size 2. Third in the sequence is our wall of height 7 and largest clique size 3.

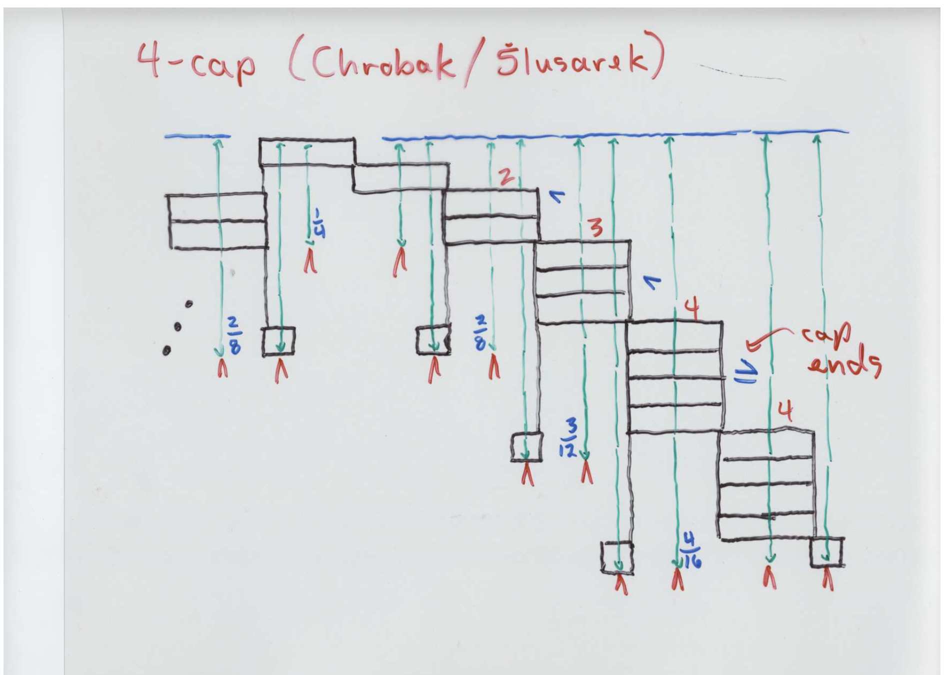
Slide 7



The sequence of walls (from the previous slide) approaches 3 in the ratio χ_{FF}/χ. This is done by attaching beneath each of 4 intervals a copy of the previous wall in the sequence. Each of those 4 intervals therefore meets (i.e., is supported by) some other in every level below. In some cases the support is local, which is to say by another of the 4; and in other cases the support is distant, which is to say by an interval in a wall of prior generation.

The structure depicted here is a *cap,* which is essentially a wall with cones standing for copies of older walls. From the top of each cone to the top of the cap, at most 1 level in 3 is occupied by an interval. This makes the cap suitable for producing a sequence of walls tending to 3 in the ratio χ_{FF}/χ. Therefore it is called a 3-*cap.* Now our search turns from walls to caps.

Slide 8

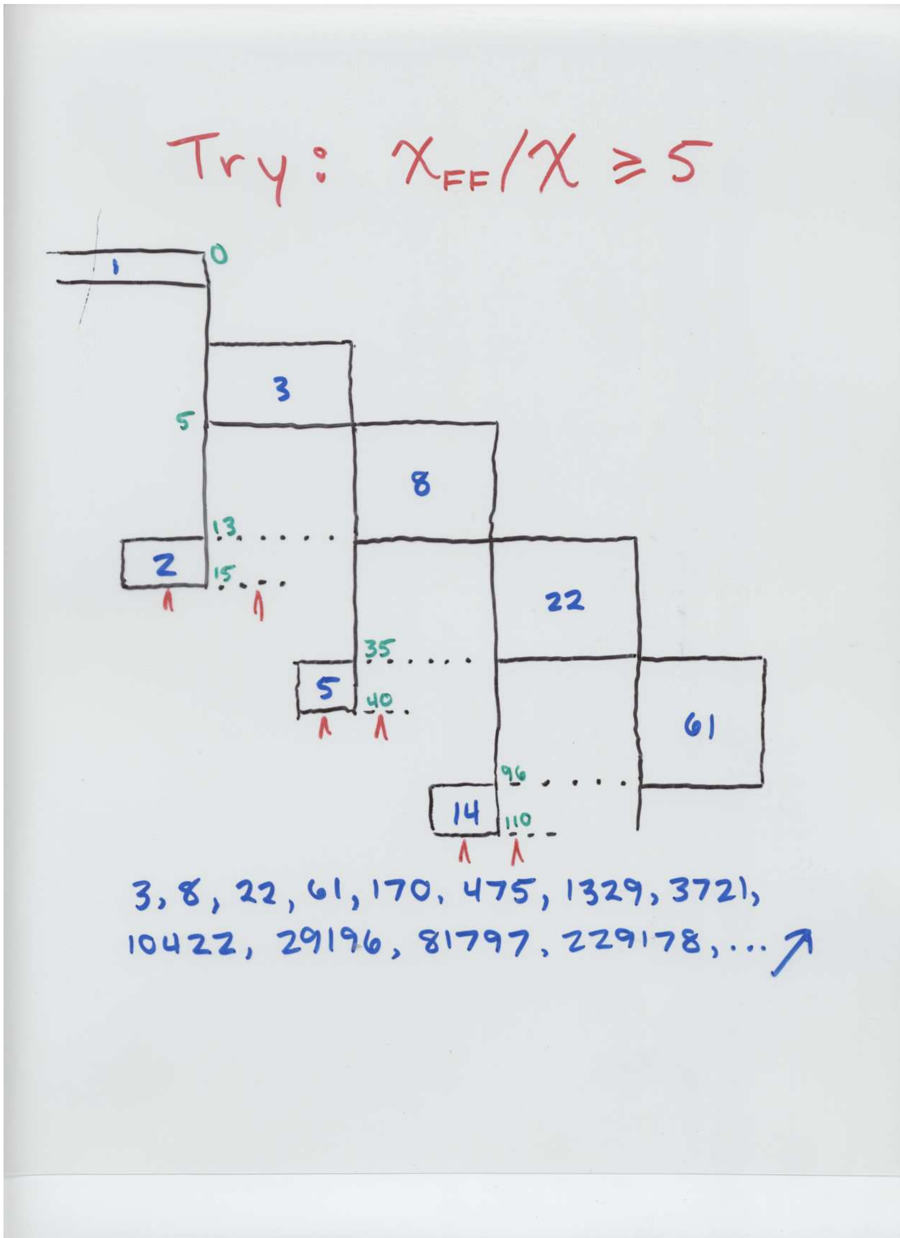


Here is a 4-cap, where intervals now are represented by unit-height blocks. At the top are twin unit blocks. (A cap should be nonempty.) Each can be supported in the middle by a cone (representing an older wall). These cone tops must be at least 4 units deeper than the top of the cap, so that above them at most 1 level in 4 is occupied by an interval. No more distant support is available for the twin unit blocks. There remains a gap of 2 levels where local support is needed, so dual blocks are added to the cap under each unit block. Now the twin unit blocks are supported, but the new block pairs at their sides still need support. (The cap has left-right symmetry. There will be no further mention of the left side.)

A cone goes beneath the 2 blocks at 8 units below cap top. Left of (and below) the 2 blocks (where there is only 1 unit block above), it is possible to introduce a sole left supporter plus another cone at the same depth. (Among 8 levels spanned by the green arrow above its cone, only 2 levels are occupied by intervals.) Now there is a need for 3 right supporters, and the 2 are properly supported. It remains to support the 3.

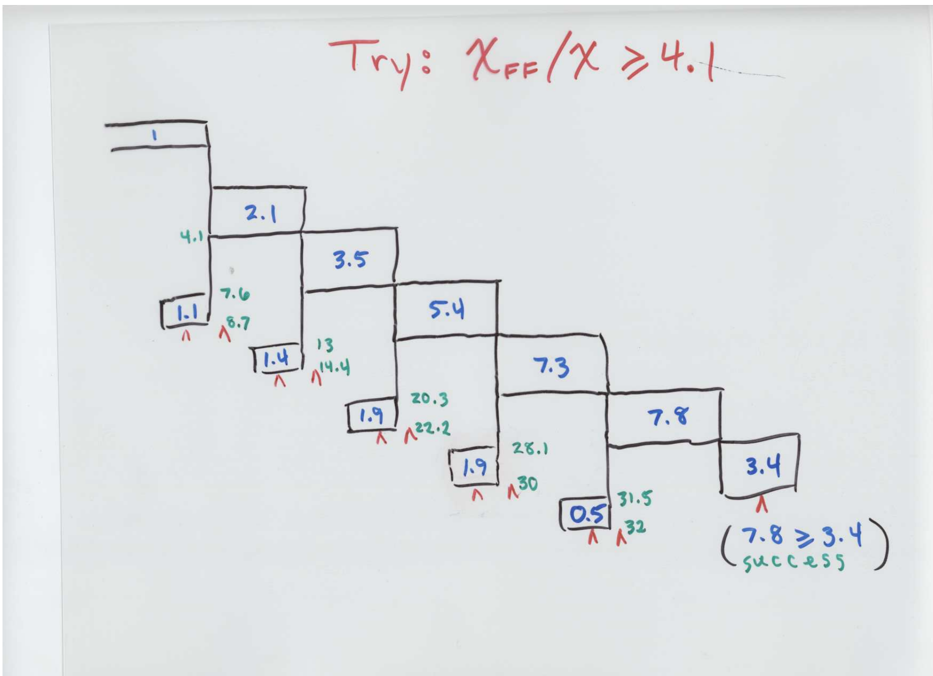
The 3 require 4 right supporters with a cone depth of 16, and those 4 get a sole left supporter as well as something new: only 4 right supporters. So the cone beneath them is at depth 16, resulting in a smaller gap than before. For right support it suffices to place another copy of what supported the first group of 4: a sole supporter at depth 16. The construction ends because the sequence of right supporters has an entry that does not exceed its predecessor.

Slide 9



After discovering a 4-cap, surely Witsenhausen, and later Chrobak and Ślusarek, must have immediately sought a 5-cap in the same way. (Here is another change in notation: a group of unit blocks is replaced by a single block with a number in the center.) It turns out that the sequence of right supporters increases strictly. Consequently, there is no 5-cap like this. A cap must be finite.

Slide 10



So one might try a ratio between 4 and 5, say 4.1, with block multiplicities rational instead of integral. Of course, the number of vertices in a cap must be an integer. This condition is easily met, as the numbers can be scaled (by some common factor) up to integers. Let us keep in mind the notion of scaling. It is important later in the construction.

Slide 11

Ślusarek obtained

$$\chi_{FF}/\chi \geqslant 4.45$$

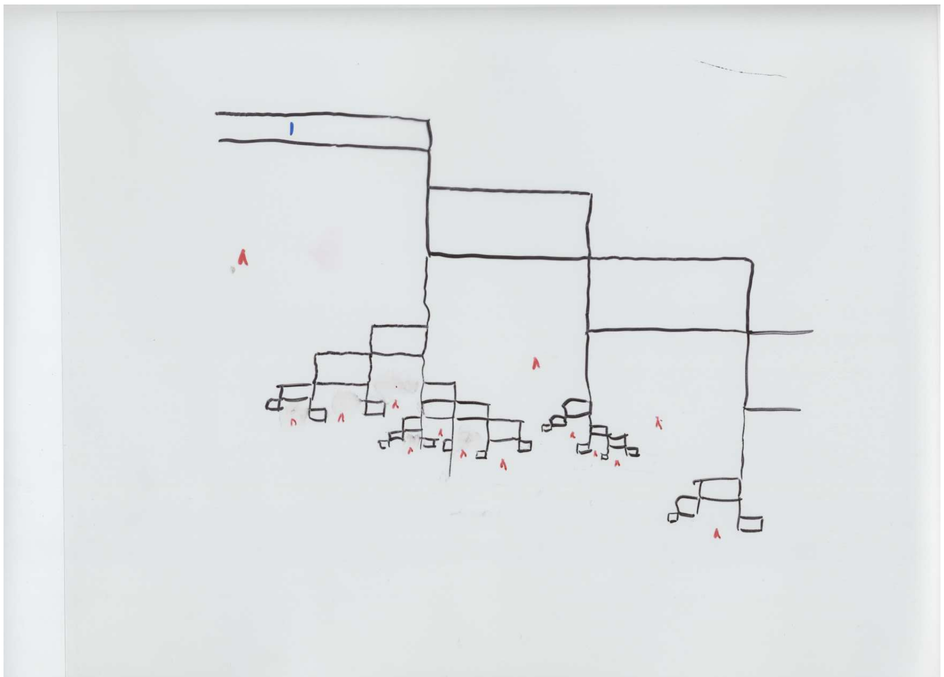
by a method like this.
He suspected 4.5
could be obtained. But
this method is exhausted at

$$1.5 + 0.5\sqrt{13+16\sqrt{2}}$$

$$\approx 4.48.$$

This method works only up to a target ratio of 4.48 or so.

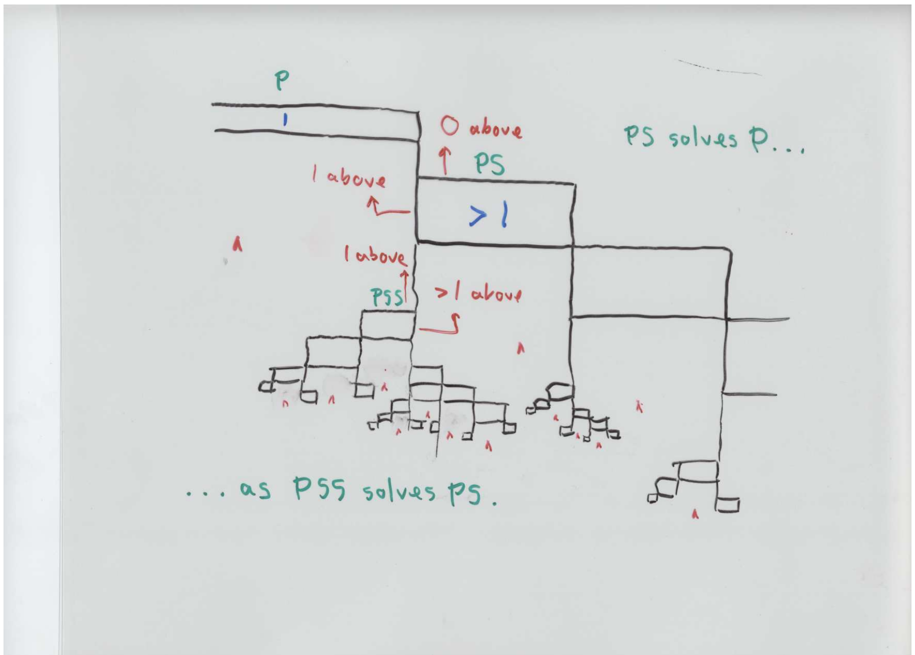
Slide 12



Some new idea is needed in order to reach a lower bound of 5. Kierstead and Trotter proposed a fuller collection of supporting blocks in the cap. Why should each left supporter be terminated immediately by a cone? Why can't each block in the cap potentially have a left supporter and a right one?

This idea turns out to be helpful, although the construction is more difficult. Each block must be supported down to the top of its cone, which is to say that a gap (of levels) must be covered jointly by left and right supporters. How should this gap be distributed between supporters so that ultimately the cap contains a finite number of blocks? Also, which of the two supporters of a given block should cover the higher levels?

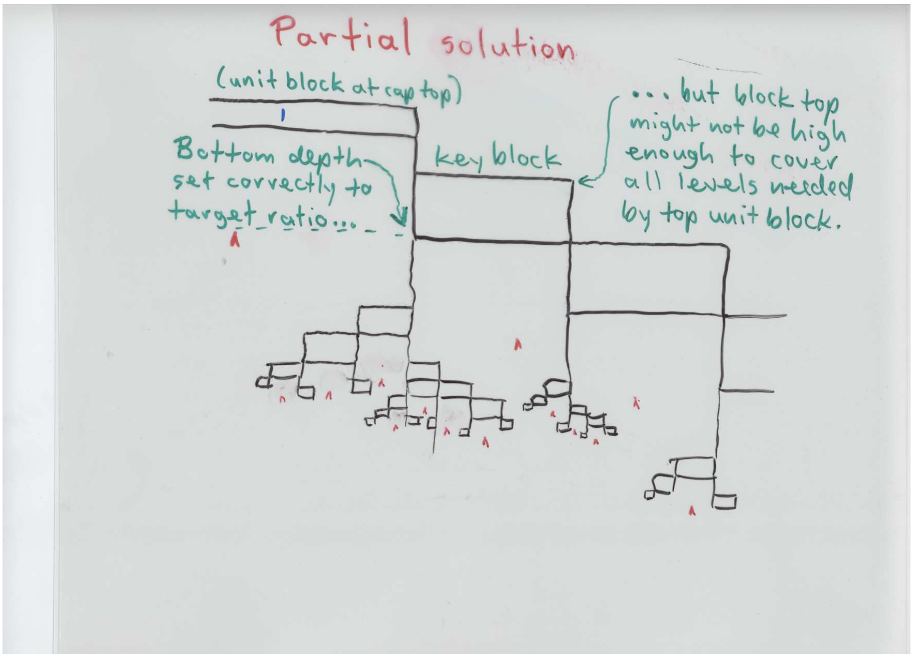
Slide 13



The present construction exploits the appearance in the cap of another kind of self-similarity. Let P be a unit block at the top of the cap. Then its right supporter PS (in the low position) along with all its descendants can be regarded as solving the problem of supporting P, subject to a certain constraint: some of those descendants have above them block P, which enlarges by 1 all the cliques they represent; while block PS itself and its other supporters do not have P above them.

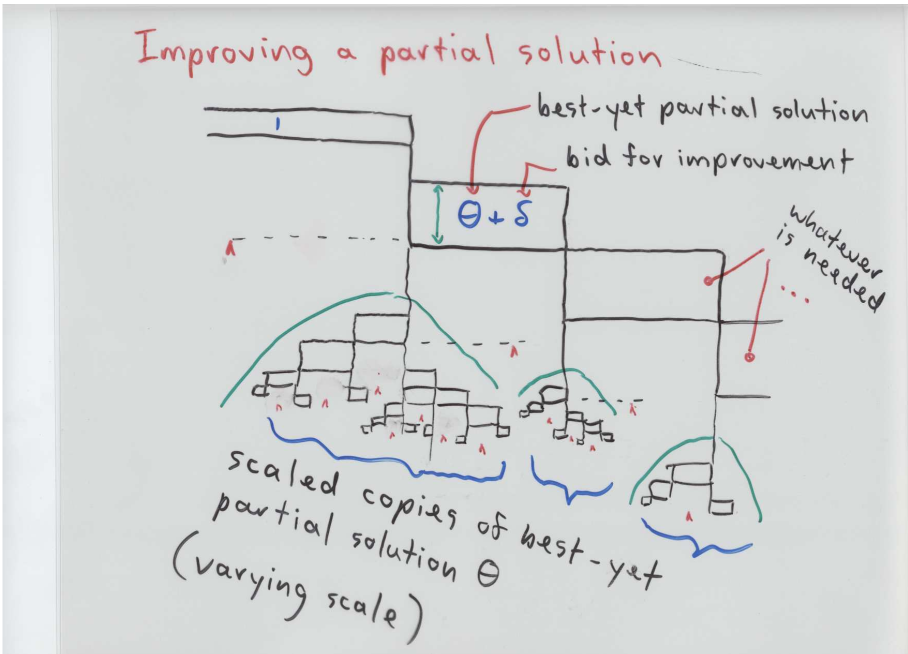
Similarly, the left supporter PSS of PS along with its descendants solves the problem of supporting PS, subject to the constraint that PSS itself and some of its supporters have only 1 unit above, while the others have PS above, which is larger than 1 (when, as in our case, the target ratio exceeds 4). So PSS is like PS reversed left to right.

Slide 14



We wish to propagate partial solutions from the PS position (here labeled "key block") to the PSS position. In a partial solution, the bottom of the key block is at the top of the cone beneath block P (where it should be), but the top of the key block is allowed to be lower than is needed to cover the gap of P.

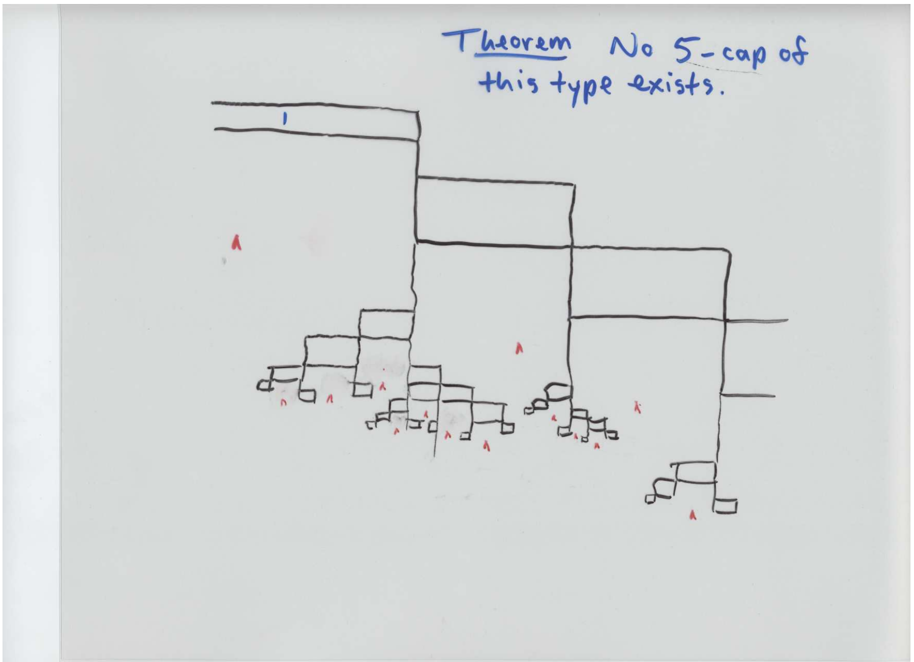
Slide 15



A partial solution will be improved by a tiny increment δ. Progress toward a complete solution is measured by the height θ of the key block in the best known partial solution. If ever θ reaches $r - 2$, where r is the target ratio (say $r = 4.9$), then the partial solution is complete, which is to say that an r-cap exists.

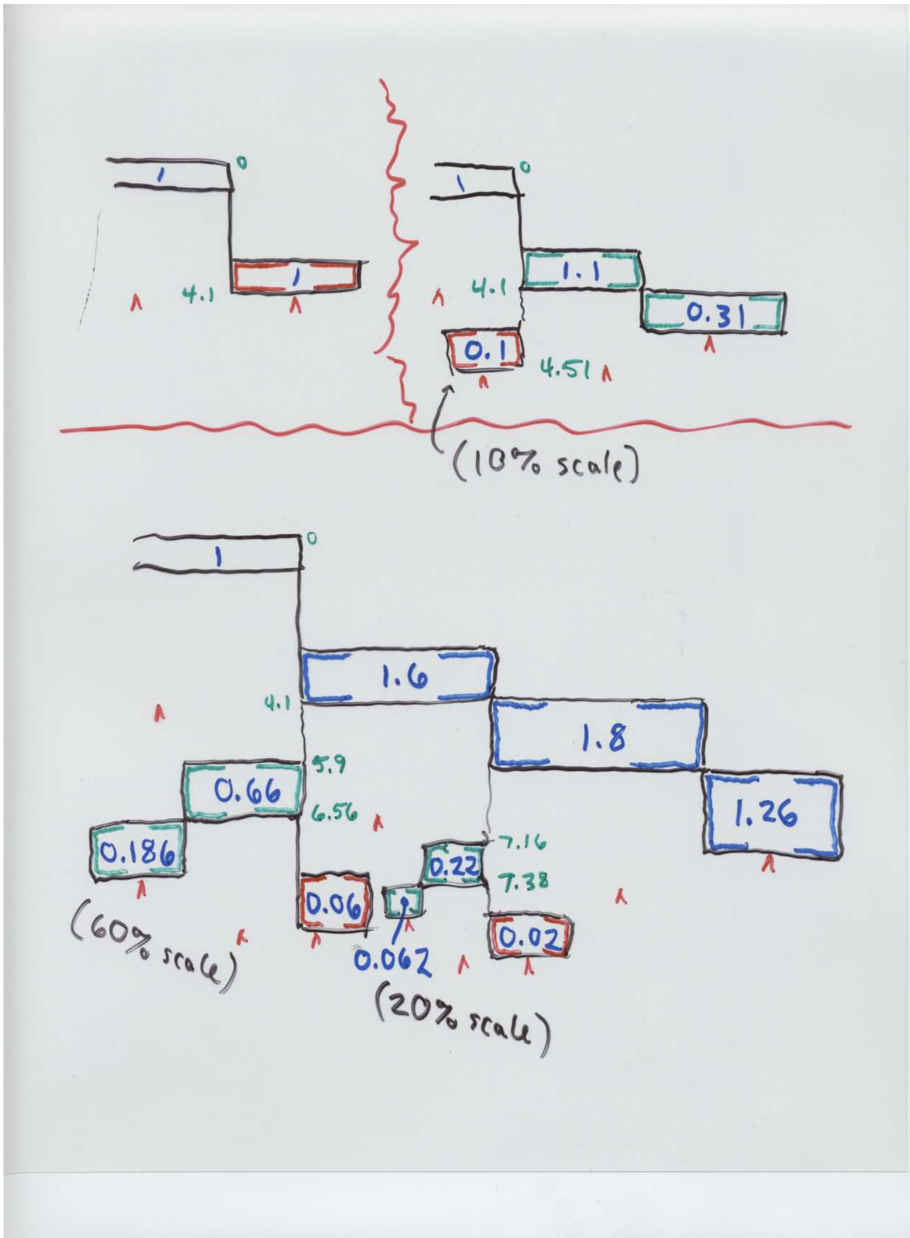
Where before there were sole left-supporting blocks immediately terminated by caps, now there are scaled copies of the best-known solution θ. The scale for each such copy is uniquely determined and may differ from scales of other copies. As before, each right supporter is determined to be whatever is needed to cover the remaining gap.

Slide 16



No 5-cap of this type exists, but we can get arbitrarily near 5.

Slide 17

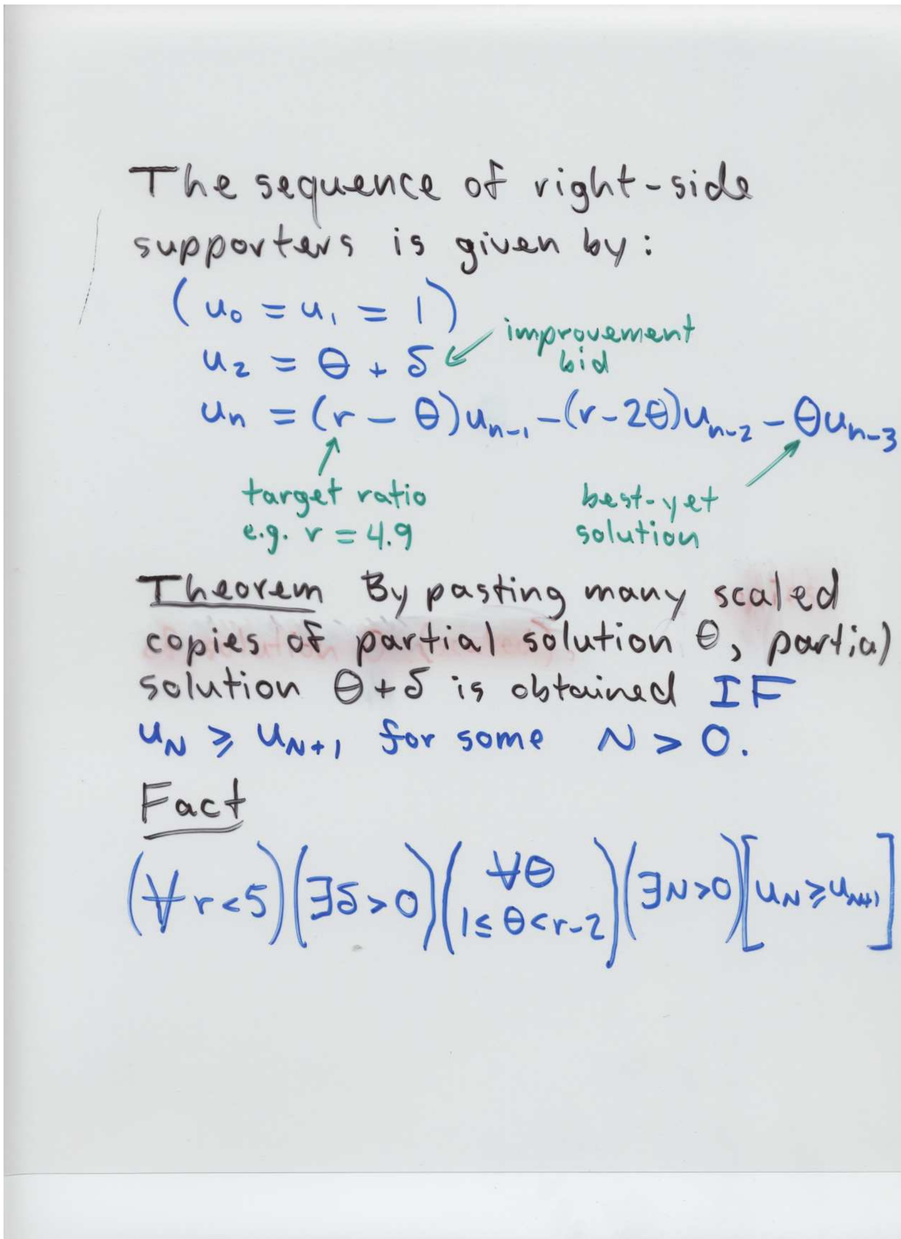


For example, here is how a 4.1-cap is obtained under the new procedure. The initial partial solution is $\theta = 1$; in this case the key block can be supported immediately by a cone. (If r were 3, then this would be a complete solution, the one with a cap of 4 intervals from many slides ago.)

We attempt an improvement of $\delta = 0.1$. (This is an arbitrary choice.) So the key block has height 1.1, and the scale of its left supporter is 10%. The remaining gap is 0.31, and such a block can be supported immediately by a cone at this depth. The bid $\delta = 0.1$ turns out to be modest enough for success.

We bid more ambitiously ($\delta = 0.5$) in the next step, and the sequence of right supporters does not end so quickly. A 60%-scaled copy of the previous solution goes under the key 1.6-block, leaving a gap of 1.8 for the right supporter. Then a 20%-scaled copy of the previous solution goes under that, leaving a gap of $1.26 \leq 1.8$. Such a block can be supported by a cone at this depth. Even $\delta = 0.5$ turns out to be modest enough, at least when $r = 4.1$ and $\theta = 1.1$. After another iteration or so, the key block reaches size $\theta = 2.1$, so a 4.1-cap exists.

Slide 18



One might wish to find a formula for the size of each block in a cap. Most important is the sequence of right supporters; when its growth halts for some $\delta > 0$, even in a single entry, some progress can be made. However, perpetual progress is not sufficient for our purpose (say when $\delta = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$).

It turns out that the sequence of right supporters satisfies a cubic linear recurrence. Analysis of this recurring sequence reveals for which r and θ one can advance from the initial partial solution ($\theta = 1$) to a complete one ($\theta \geq r - 2$). One can show that when $r < 5$, some uniform choice of $\delta > 0$ suffices to reach the goal.

Slide 19

$$u_0 = u_1 = 1$$
$$u_2 = \theta + \delta$$
$$u_n = (r-\theta)u_{n-1} - (r-2\theta)u_{n-2} - \theta u_{n-3}$$
$$\equiv$$
$$\sum_{n \geqslant 0} u_n x^n = \frac{1 + (1+\theta-r)x + \delta x^2}{q(x)}$$
$$\sum_{n \geqslant 0} (u_{n+1} - u_n)x^n = \frac{x[(\theta+\delta)(1-x) - 1]}{q(x)}$$

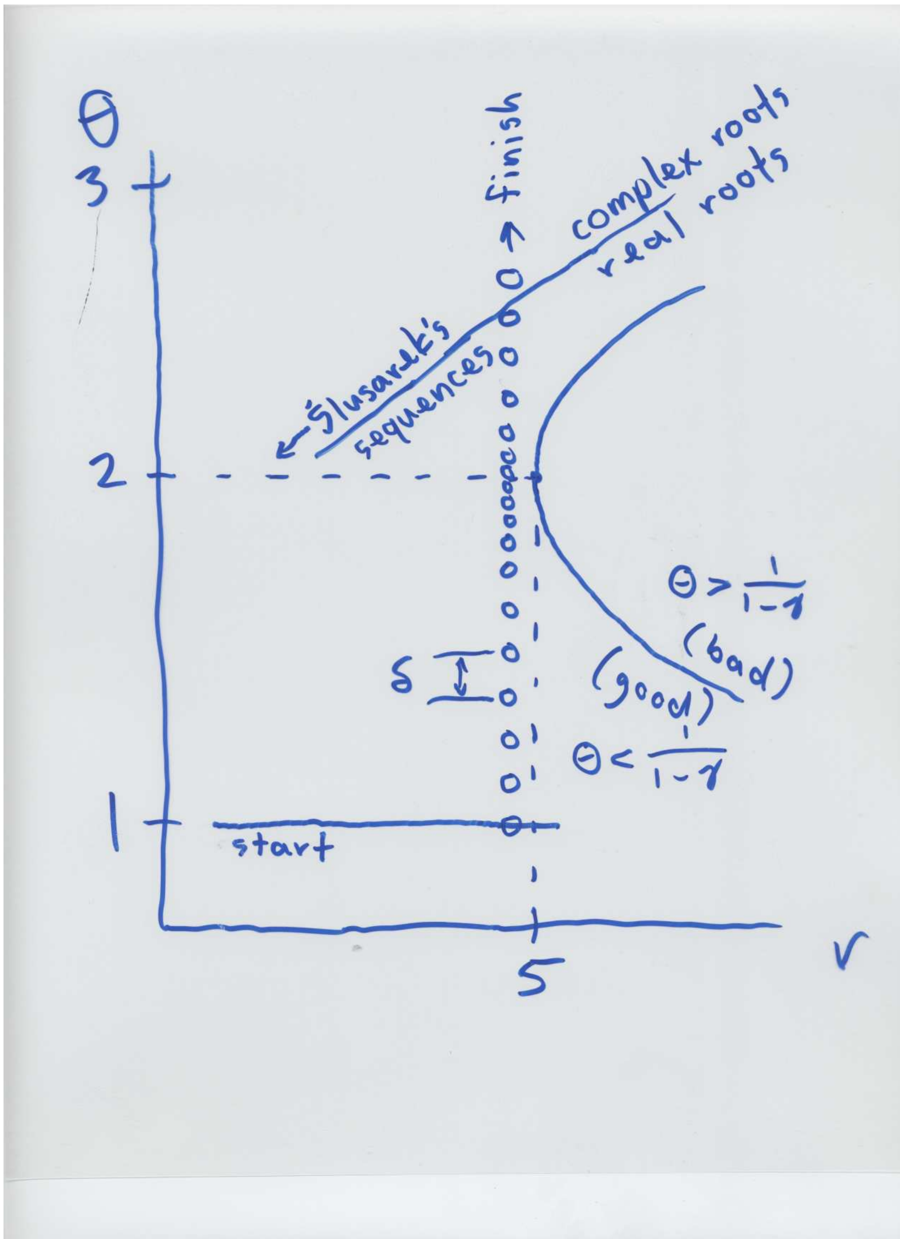
where

$$q(x) = 1 - (r-\theta)x + (r-2\theta)x^2 + \theta x^3$$
$$= 1 + rx(x-1) + \theta x(x-1)^2.$$

It is well known that a linear recurring sequence is the sequence of coefficients in the Taylor series of a rational function. Here are generating functions for the sequence of right supporters and its difference sequence, whose asymptotic behavior is crucial.

Each cap in the present construction is essentially a finite collection of scaled instances of this sequence featuring one value of r (near 5) and many values of θ ranging from 1 in the bottom (interior) of the cap to $r-2$ at the top (exterior).

Slide 20



Such a cap is represented by open circles in this slide. The depicted curve in the r, θ-plane is the boundary between sequence instances that increase strictly (i.e., are bad) and those that are good for our construction. (Assume

δ vanishes.) This is decided by comparing θ to $\frac{1}{1-\gamma}$, where γ is the smallest root of the characteristic polynomial of the recurrence relation (γ depends on r and θ, but not δ).

The r-minimum point on that curve is precisely $(r, \theta, \delta) = (5, 2, 0)$, known as the Fibonacci sequence, which is of course strictly increasing (i.e., bad) after the initial entries $1, 1$. The Fibonacci sequence arises in our proof that there is no 5-cap, as it happens. In the search for an r-cap with $r < 5$, it is natural to consider sequences near the Fibonacci sequence. Who would have thought though first to lift the recurrence from quadratic to cubic? And who would have thought that the instances nearest the Fibonacci sequence belong neither at the top nor bottom of the cap, but in the middle?