Cover’s Dichotomy Counting Theorem

(Thomas Cover, 1965)

Statement

Consider K points in general position in $\mathbb{R}^N$: $S = \{ x^{(1)}, x^{(2)}, \dots, x^{(K)} \}$.
Consider all possible dichotomies of S. That is: partitions into $S_{\oplus}$ and $S_{\ominus}$.
There are 2^K such partitions.
Any such dichotomy is either linearly separable or it is not.
Further assume that $N + 1 < K$.
Cover's Thm states that (again – for any such configuration in general positions) the number of linearly separable dichotomies is

$$CovNum(K, N) = 2 \sum_{i=0}^{N} \binom{K-1}{i}$$

For $K \leq N + 1$ all 2^K dichotomies are linearly separable.

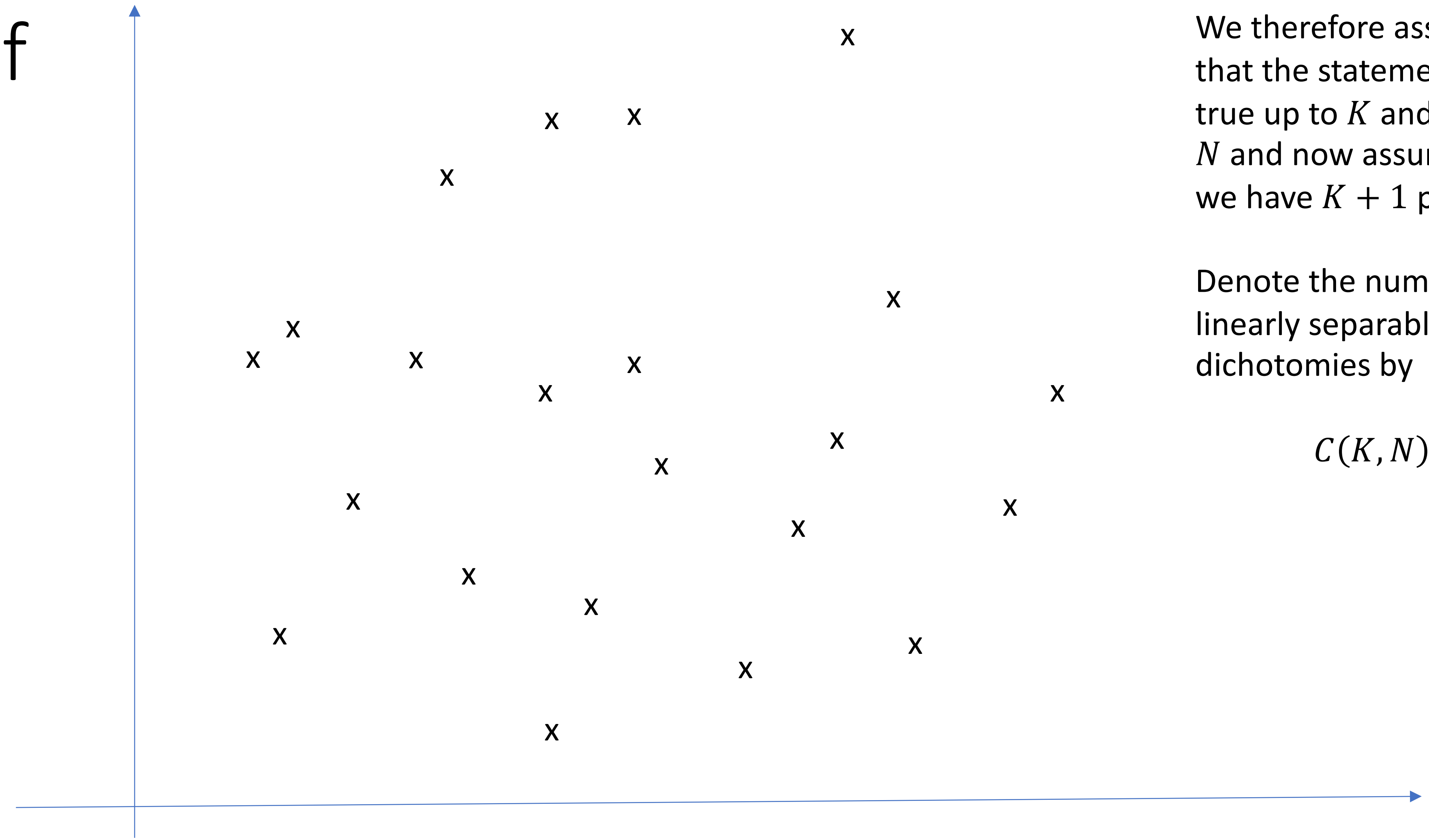
Comments

- General position means that, for all $d \leq N$ there are no d+1 points that reside on the same $d-1$ dimensional hyperplane. For example – no 2 points are identical, no 3 points are co-linear, no 4 points are on the same 2 dimensional hyperplane, etc …

- A immediate conclusion from the theorem is as follows. If we draw a dichotomy, uniformly at random, for a set S as above, then the probability that this dichotomy is linearly separable is (exactly!):

$$\frac{1}{2^{K-1}} \sum_{i=0}^{N} \binom{K-1}{i}$$

Proof

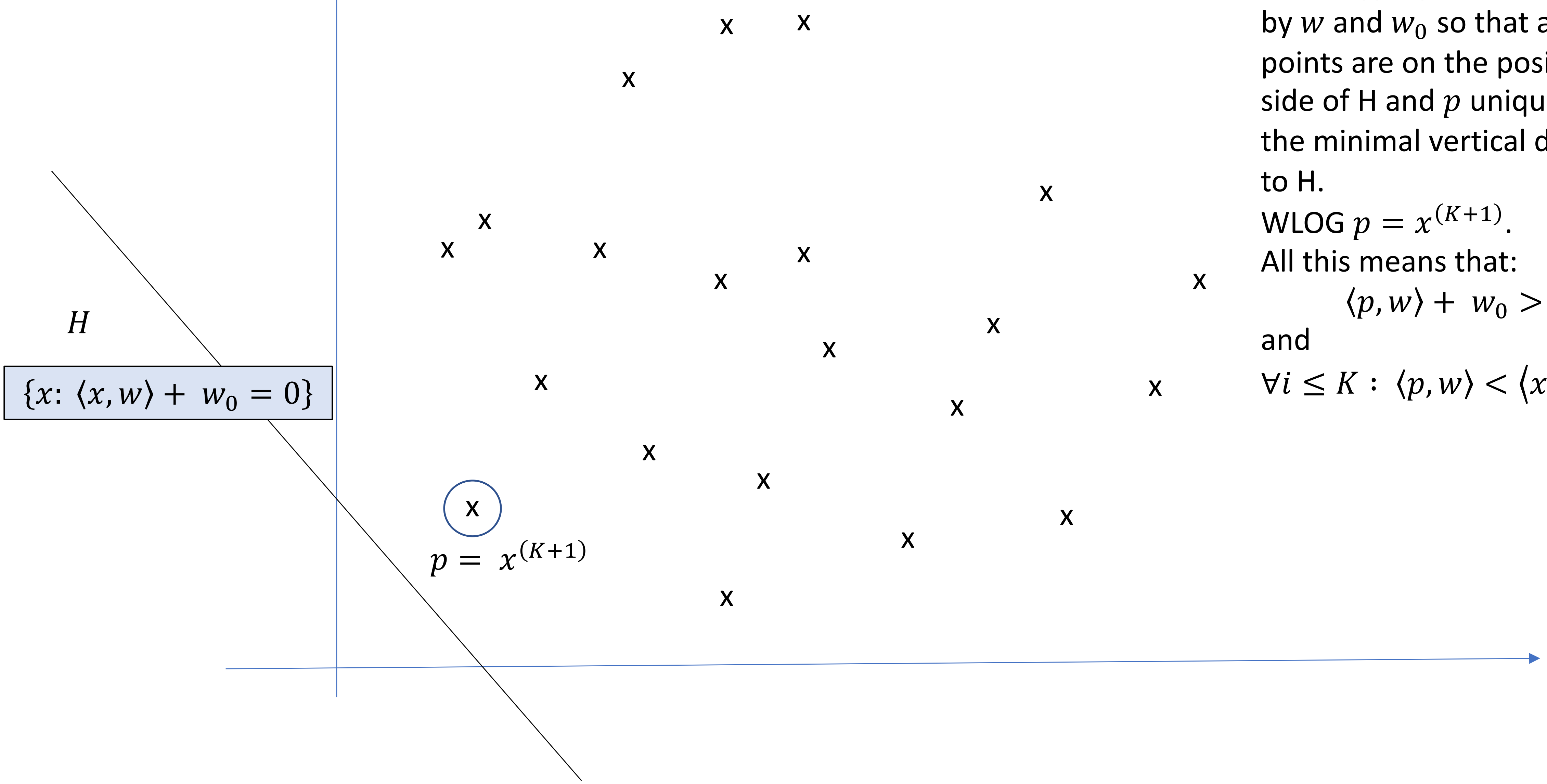


We will use induction. We therefore assume that the statement is true up to K and for all N and now assume that we have $K + 1$ points.

Denote the number of linearly separable dichotomies by

$$C(K, N)$$

Proof



Find a point $p = x^{(i)} \in S$ and a hyperplane H, defined by w and w_0 so that all points are on the positive side of H and p uniquely has the minimal vertical distance to H.

WLOG $p = x^{(K+1)}$.

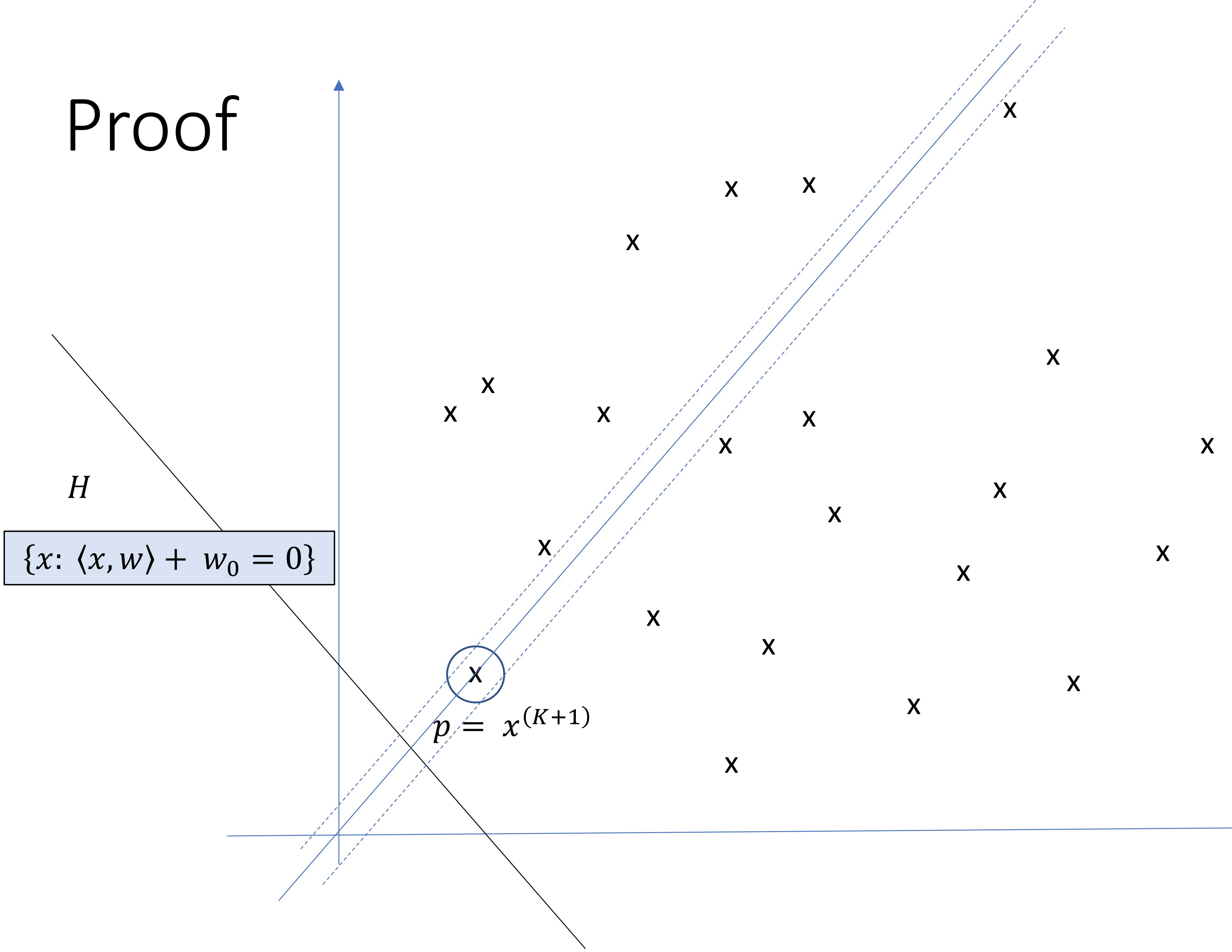
All this means that:

$$\langle p, w\rangle + w_0 > 0$$

and

$$\forall i \leq K : \langle p, w\rangle < \langle x^{(i)}, w\rangle$$

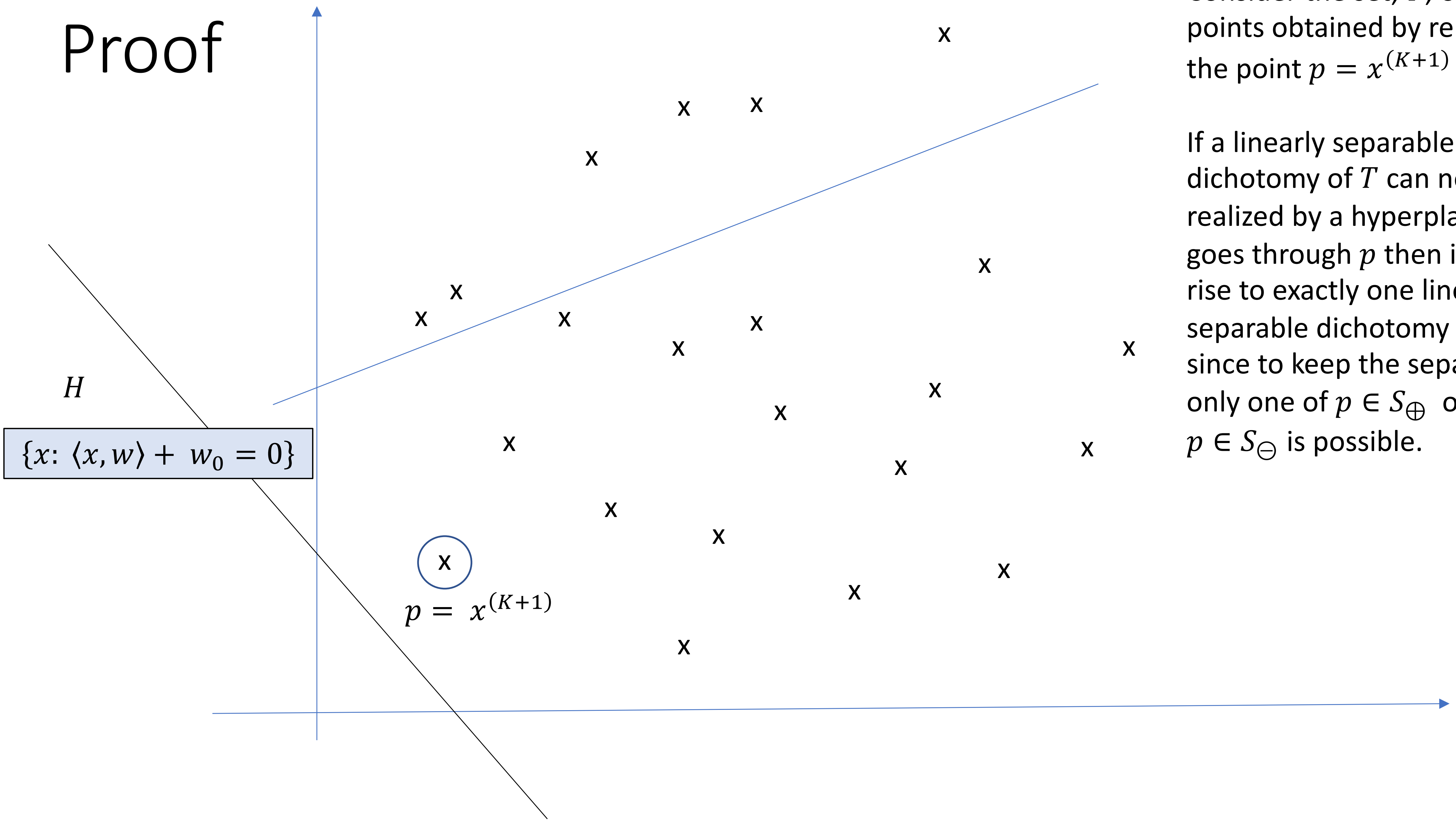
Proof



Consider the set, T, of K points obtained by removing the point $p = x^{(K+1)}$ from S.

If a linearly separable dichotomy of T can be realized by a hyperplane that goes through p then it gives rise to exactly two linearly separable dichotomies of S, by moving the hyperplane to either side and setting $p \in S_{\oplus}$ or $p \in S_{\ominus}$, accordingly.

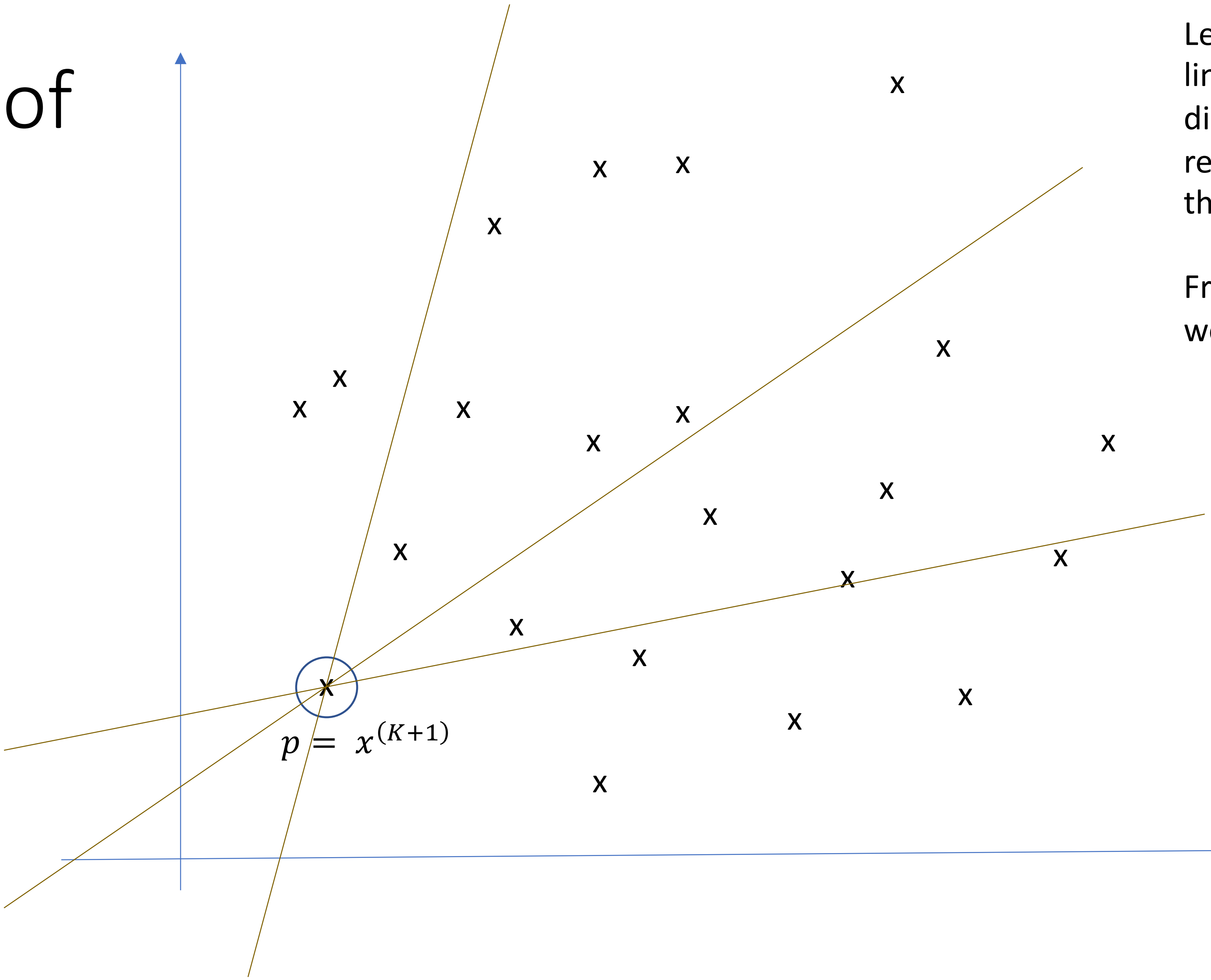
Proof



Consider the set, T, of K points obtained by removing the point $p = x^{(K+1)}$ from S.

If a linearly separable dichotomy of T can not be realized by a hyperplane that goes through p then it gives rise to exactly one linearly separable dichotomy of S, since to keep the separability only one of $p \in S_{\oplus}$ or $p \in S_{\ominus}$ is possible.

Proof

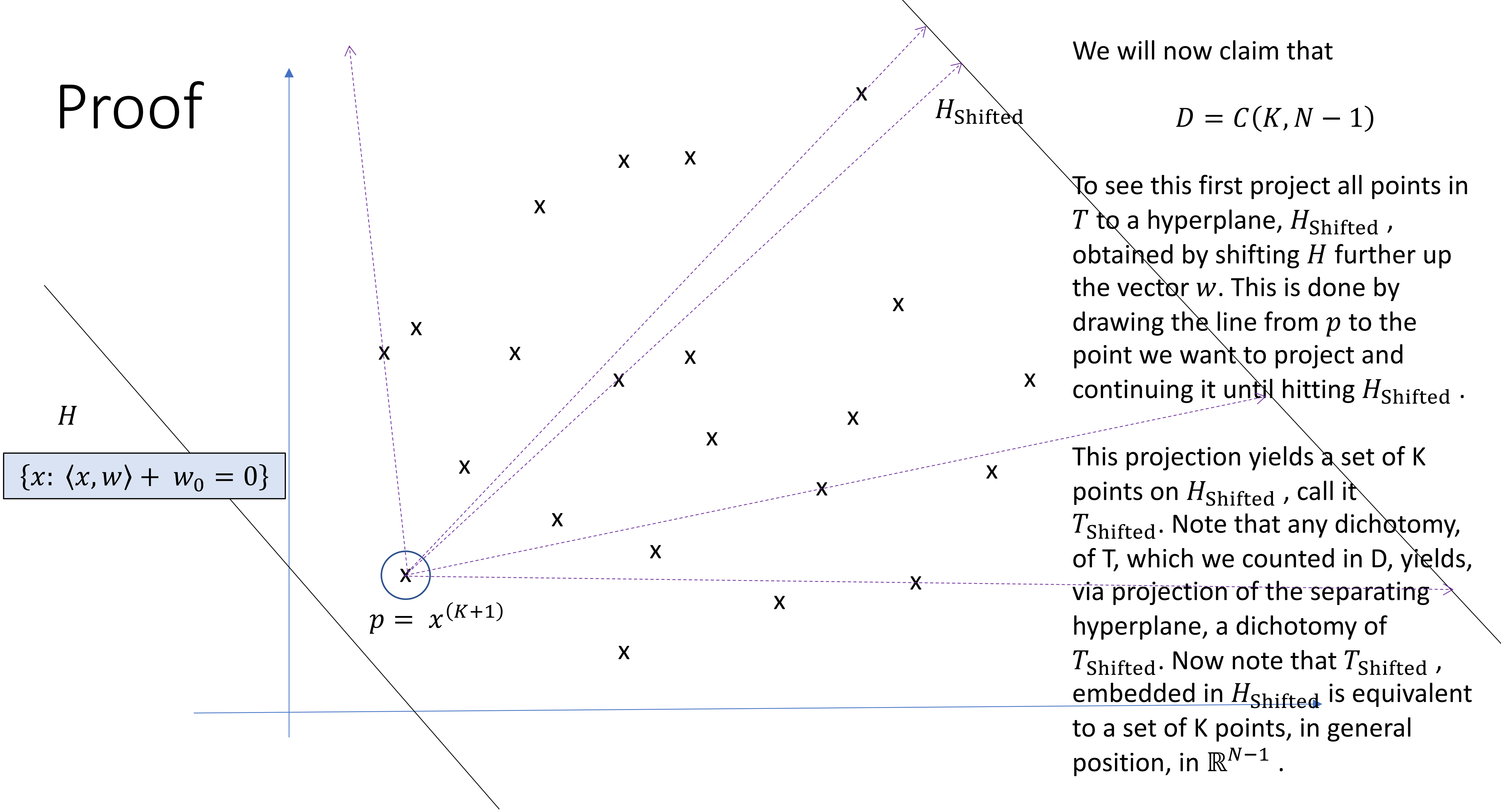


Let D denote the number of linearly separable dichotomies of T that can be realized with a hyperplane through p.

From the previous two slides we get:

$$C(K+1, N)$$
$$= C(K,N) - D + 2D$$
$$= C(K,N) + D$$

Proof



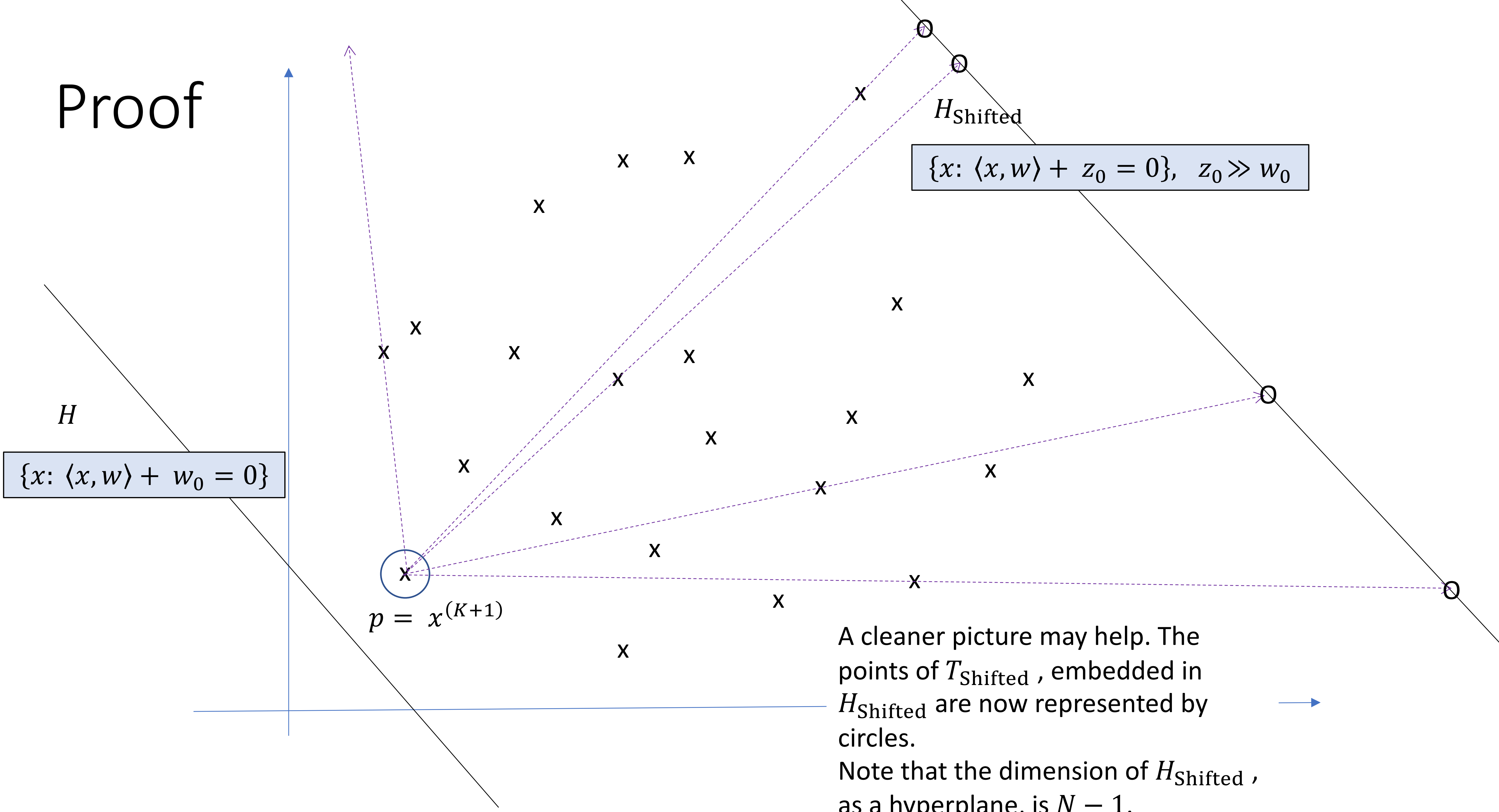
We will now claim that

$$D = C(K, N-1)$$

To see this first project all points in T to a hyperplane, H_{Shifted}, obtained by shifting H further up the vector w. This is done by drawing the line from p to the point we want to project and continuing it until hitting H_{Shifted}.

This projection yields a set of K points on H_{Shifted}, call it T_{Shifted}. Note that any dichotomy, of T, which we counted in D, yields, via projection of the separating hyperplane, a dichotomy of T_{Shifted}. Now note that T_{Shifted}, embedded in H_{Shifted} is equivalent to a set of K points, in general position, in $\mathbb{R}^{N-1}$.

Proof



A cleaner picture may help. The points of T_{Shifted} , embedded in H_{Shifted} are now represented by circles.

Note that the dimension of H_{Shifted} , as a hyperplane, is $N-1$.

Proof – final step, some algebra …

We now have the recurrence relation:

$$\begin{aligned}&C(K+1,N)\\&= C(K,N) - D + 2D\\&= C(K,N) + D\\&= C(K,N) + C(K,N-1)\end{aligned}$$

By induction we then have:

$$C(K+1,N) = 2\sum_{i=0}^{N}\binom{K-1}{i} + 2\sum_{i=0}^{N-1}\binom{K-1}{i} = 2\sum_{i=0}^{N}\binom{K}{i}$$

Which completes our inductive step.

The last equality uses the identity $\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$ and some shifting of the summation.

$\boxed{\mathcal{QED}}$