In short, we have proved properties of LPVC(G). There exists a half-integral optimal solution $(x_v)_{v\in V(G)}$ to LPVC(G), and it can be found efficiently. We can look at this solution as a partition of $V(G)$ into parts V_0, $V_{\frac{1}{2}}$, and V_1 with the following message: greedily take V_1 into a solution, do not take any vertex of V_0 into a solution, and in $V_{\frac{1}{2}}$, we do not know what to do and that is the hard part of the problem. However, as an optimum solution pays $\frac{1}{2}$ for every vertex of $V_{\frac{1}{2}}$, the hard part — the kernel of the problem — cannot have more than $2k$ vertices.

2.6 Sunflower lemma

In this section we introduce a classical result of Erdős and Rado and show some of its applications in kernelization. In the literature it is known as the sunflower lemma or as the Erdős-Rado lemma. We first define the terminology used in the statement of the lemma. A *sunflower* with k *petals* and a *core* Y is a collection of sets $S_1, \dots, S_k$ such that $S_i \cap S_j = Y$ for all $i \neq j$; the sets $S_i \setminus Y$ are petals and we require *none of them to be empty.* Note that a family of pairwise disjoint sets is a sunflower (with an empty core).

Theorem 2.25 (Sunflower lemma). *Let $\mathcal{A}$ be a family of sets (without duplicates) over a universe U, such that each set in $\mathcal{A}$ has cardinality exactly d. If $|\mathcal{A}| > d!(k-1)^d$, then $\mathcal{A}$ contains a sunflower with k petals and such a sunflower can be computed in time polynomial in $|\mathcal{A}|$, $|U|$, and k.*

Proof. We prove the theorem by induction on d. For $d = 1$, i.e., for a family of singletons, the statement trivially holds. Let $d \geq 2$ and let $\mathcal{A}$ be a family of sets of cardinality at most d over a universe U such that $|\mathcal{A}| > d!(k-1)^d$.

Let $\mathcal{G} = \{S_1, \dots, S_\ell\} \subseteq \mathcal{A}$ be an inclusion-wise maximal family of pairwise disjoint sets in $\mathcal{A}$. If $\ell \geq k$ then $\mathcal{G}$ is a sunflower with at least k petals. Thus we assume that $\ell < k$. Let $S = \bigcup_{i=1}^{\ell} S_i$. Then $|S| \leq d(k-1)$. Because $\mathcal{G}$ is maximal, every set $A \in \mathcal{A}$ intersects at least one set from $\mathcal{G}$, i.e., $A \cap S \neq \emptyset$. Therefore, there is an element $u \in U$ contained in at least

$$\frac{|\mathcal{A}|}{|S|} > \frac{d!(k-1)^d}{d(k-1)} = (d-1)!(k-1)^{d-1}$$

sets from $\mathcal{A}$. We take all sets of $\mathcal{A}$ containing such an element u, and construct a family $\mathcal{A}'$ of sets of cardinality $d-1$ by removing from each set the element u. Because $|\mathcal{A}'| > (d-1)!(k-1)^{d-1}$, by the induction hypothesis, $\mathcal{A}'$ contains a sunflower $\{S'_1, \dots, S'_k\}$ with k petals. Then $\{S'_1 \cup \{u\}, \dots, S'_k \cup \{u\}\}$ is a sunflower in $\mathcal{A}$ with k petals.

The proof can be easily transformed into a polynomial-time algorithm, as follows. Greedily select a maximal set of pairwise disjoint sets. If the size of this set is at least k, then return this set. Otherwise, find an element u contained in the maximum number of sets in $\mathcal{A}$, and call the algorithm recursively on sets of cardinality $d-1$, obtained from deleting u from the sets containing u. $\square$

2.6.1 *d*-Hitting Set

As an application of the sunflower lemma, we give a kernel for d-Hitting Set. In this problem, we are given a family $\mathcal{A}$ of sets over a universe U, where each set in the family has cardinality at most d, and a positive integer k. The objective is to decide whether there is a subset $H \subseteq U$ of size at most k such that H contains at least one element from each set in $\mathcal{A}$.

Theorem 2.26. d-Hitting Set *admits a kernel with at most* $d!k^d$ *sets and at most* $d!k^d \cdot d^2$ *elements.*

Proof. The crucial observation is that if $\mathcal{A}$ contains a sunflower

$$S = \{S_1, \dots, S_{k+1}\}$$

of cardinality $k+1$, then every hitting set H of $\mathcal{A}$ of cardinality at most k intersects the core Y of the sunflower S. Indeed, if H does not intersect Y, it should intersect each of the $k+1$ disjoint petals $S_i \setminus Y$. This leads to the following reduction rule.

Reduction HS.1. Let $(U, \mathcal{A}, k)$ be an instance of d-Hitting Set and assume that $\mathcal{A}$ contains a sunflower $S = \{S_1, \dots, S_{k+1}\}$ of cardinality $k+1$ with core Y. Then return $(U', \mathcal{A}', k)$, where $\mathcal{A}' = (\mathcal{A} \setminus S) \cup \{Y\}$ is obtained from $\mathcal{A}$ by deleting all sets $\{S_1, \dots, S_{k+1}\}$ and by adding a new set Y and $U' = \bigcup_{X \in \mathcal{A}'} X$.

Note that when deleting sets we do not delete the elements contained in these sets but only those which do not belong to any set. Then the instances $(U, \mathcal{A}, k)$ and $(U', \mathcal{A}', k)$ are equivalent, i.e. $(U, \mathcal{A})$ contains a hitting set of size k if and only if $(U, \mathcal{A}')$ does.

The kernelization algorithm is as follows. If for some $d' \in \{1, \dots, d\}$ the number of sets in $\mathcal{A}$ of size exactly d' is more than $d'!k^{d'}$, then the kernelization algorithm applies the sunflower lemma to find a sunflower of size $k+1$, and applies Reduction HS.1 on this sunflower. It applies this procedure exhaustively, and obtains a new family of sets $\mathcal{A}'$ of size at most $d!k^d \cdot d$. If $\emptyset \in \mathcal{A}'$ (that is, at some point a sunflower with an empty core has been discovered), then the algorithm concludes that there is no hitting set of size at most k and returns that the given instance is a no-instance. Otherwise, every set contains at most d elements, and thus the number of elements in the kernel is at most $d!k^d \cdot d^2$. $\square$