A Guided Tour of Mathematical Optimization

Levent Tunçel

Dept. of Combinatorics and Optimization,

Faculty of Mathematics,

University of Waterloo,

Canada.

`ltuncel@math.uwaterloo.ca`

November 27, 2010

1 Linear Optimization

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ be given. Then, we have the LP problem

$$
\begin{array}{llrcl}
(LP_1) & \text{Maximize} & c^T x & & \\
 & & Ax & \leq & b, \\
 & & (x & \in & \mathbb{R}^n).
\end{array}
$$

All vectors are column vectors.

$$u, v \in \mathbb{R}^m, u \leq v \text{ means } u_i \leq v_i, \forall i \in \{1, 2, \dots, m\}.$$

Theorem 1.1 *(Fundamental Theorem of LP) For every LP, exactly one of the following is true:*

- *LP has no solution (LP is infeasible)*
- *LP is unbounded*
  ($\forall M \in \mathbb{R}$, $\exists x \in \mathbb{R}^n$ such that $Ax \leq b$ and $c^T x > M$)
- *LP has optimal solution(s)*
  ($\exists \bar{x} \in \mathbb{R}^n$ such that $A\bar{x} \leq b$ and $\forall x \in \mathbb{R}^n$ satisfying $Ax \leq b$, $c^T \bar{x} \geq c^T x$).

Note:

$$Ax \leq b$$

iff

$$Ax + s = b, s \geq 0, \text{ for some } s \in \mathbb{R}^m$$

iff

$$Au - Av + s = b, u \geq 0, v \geq 0, s \geq 0, \text{ for some } u, v \in \mathbb{R}^n, s \in \mathbb{R}^m.$$

The last system has the form

$$\tilde{A}\tilde{x} = b, \tilde{x} \geq 0.$$

Lemma 1.1 *(Farkas' Lemma) Let* $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ *be given. Then exactly one of the following systems has a solution:*

(I) $Ax = b, x \geq 0$;

(II) $A^T y \geq 0, b^T y < 0.$

(Farkas' Lemma) Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ be given. Then exactly one of the following systems has a solution:

(I) $Ax = b, x \geq 0;$

(II) $A^T y \geq 0, b^T y < 0.$

Recall the Fundamental Theorem of Linear Algebra:

Theorem 1.2 *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ be given. Then exactly one of the following systems has a solution:*

(I) $Ax = b$*;*

(II) $A^T y = 0, b^T y < 0.$

Farkas' Lemma is a powerful generalization of the Fundamental Theorem of Linear Algebra.

$$
\begin{array}{llrcl}
(LP_1) & \text{Maximize} & c^T x & & \\
 & & Ax & \leq & b, \\
 & & (x & \in & \mathbb{R}^n).
\end{array}
$$

Suppose (LP_1) has an optimal solution. Let z^* denote the optimal objective value. If we have $\bar{x} \in \mathbb{R}^n$ such that $A\bar{x} \leq b$ then $c^T\bar{x} \leq z^*$.

If we do not know z^*, how do we prove

$c^T\bar{x} \geq z^* - 10^{-8}$ or even $c^T\bar{x} \geq z^* - 10^8$?

Therefore, we define the dual of (LP_1):

$$
\begin{array}{llrcl}
(LD_1) & \text{Minimize} & b^T u & & \\
 & & A^T u & = & c, \\
 & & u & \geq & 0.
\end{array}
$$

Let us use another form for our linear optimization problem:

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ be given. Then, we have the LP problem

$$\begin{array}{lllll} (LP) & \text{Max} & c^T x & & \\ & & Ax & = & b, \\ & & x & \geq & 0. \end{array}$$

The *dual* of (LP) is defined (or directly derived from the definition of the dual above) as

$$
\begin{array}{llrcl}
(LD) & \text{Min} & b^T y & & \\
 & & A^T y & \geq & c.
\end{array}
$$

Theorem 1.3 *Let A, b, c be given as above, defining (LP) and (LD). Then*

(a) (Weak Duality Relation) For every $\bar{x}$ feasible in (LP), for every $\bar{y}$ feasible in (LD), $c^T\bar{x} \leq b^T\bar{y}$; moreover, if in addition, $c^T\bar{x} = b^T\bar{y}$ then $\bar{x}$ optimal in (LP) and $\bar{y}$ is optimal in (LD);

(b) (Duality Theorem, type I) if (LP) has an optimal solution, then so does (LD) and their optimal objective values coincide;

(c) (Duality Theorem, type II) if (LP) and (LD) both have feasible solutions, then they both have optimal solutions and their optimal objective values are the same.

Lemma 1.2 *Suppose* (LP) *has a feasible solution. Then the* (LP) *has a feasible solution* $\bar{x} \in \mathbb{R}^n$ *such that*

$$|\{j : \bar{x}_j \neq 0\}| \leq m.$$

Proof: Let $x \in \mathbb{R}^n$ be a feasible solution of (LP). Repeat the following:

- $B := \{j : \ x_j > 0\}, \ \ N := \{1, 2, \ldots, n\} \backslash B.$
- Find a linear dependence among the columns $\{A_j : j \in B\}$. If none, then STOP.
- Otherwise, we have $d \in \mathbb{R}^n$ such that $Ad = 0$, $d_N = 0$, $d_B \neq 0$.
- Choose the sign of d such that d_B has a negative component.
- Let $\alpha := \max\{\alpha > 0 : x + \alpha d \geq 0\}$.
- Set $x := x + \alpha d$.

■

Note that in the above algorithm, the number of arithmetic operations is $O(n^2$.(the number of positive components of original x)).

Theorem 1.4 *(Helly's Theorem for halfspaces and polyhedra) A system of $n \geq (m+1)$ linear inequalities in m variables has a solution iff every subsystem of $(m+1)$ of the linear inequalities has a solution.*

Proof: Let us represent the given system as

$$(I) \qquad A^T y \geq c,$$

where $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$ are given such that $n \geq (m+1)$. If the system (I) has a solution, every subsystem of it must have a solution.

Suppose (I) has no solution. Consider the pair of primal-dual LP problems:

$$(P) \quad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax = 0 \\ & x \geq 0 \end{array} \qquad \text{and} \qquad (D) \quad \begin{array}{ll} \text{minimize} & 0^T y \\ \text{subject to} & A^T y \geq c. \end{array}$$

Note that (P) is always feasible (consider $x := 0$). Therefore, by the duality theorem and the fundamental theorem of LP, (D) is infeasible iff (P) is unbounded.

Since, (I) has no solution, we conclude that there exists $\hat{x} \in F$, where

$$F := \left\{x \in \mathbb{R}^n : Ax = 0, c^T x = 1, x \geq 0\right\}.$$

By the last lemma, there exists $\bar{x} \in F$ such that

$$|\{j : \bar{x}_j \neq 0\}| \leq (m+1).$$

Let $J := \{j : \bar{x}_j > 0\}$. Then we have that the system

$$A_J x_J = 0, c_J^T x_J = 1, x_J \geq 0$$

has a solution. Therefore, (by the easy part of Farkas' Lemma), the alternative system

$$A_J^T y + \eta c_J \geq 0, \eta < 0$$

has no solution. ■

$F \subseteq \mathbb{R}^m$ is a polyhedron if $F = \{x \in \mathbb{R}^m : Ax \leq b\}$ for some $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$ for some n.

Corollary 1.1 *Let $P_1, P_2, \ldots, P_r$ be polyhedra in $\mathbb{R}^m$, $r \geq (m+1)$. Then,*

$$\bigcap_{i=1}^{r} P_i \neq \emptyset$$

iff

$$\bigcap_{i \in J} P_i \neq \emptyset, \text{ for all } J \subseteq \{1, 2, \ldots, r\} \text{ such that } |J| = m + 1.$$

The last lemma which we used in proving Helly's Theorem has close ties to Carathéodory's Theorem.

Line segment joining u and v in $\mathbb{R}^n$:

$$\{\lambda u + (1-\lambda)v : \lambda \in [0,1]\}.$$

Convex sets: $F \subseteq \mathbb{R}^n$ is convex if for every pair of points $u, v \in F$, the line segment joining u and v lies entirely in F.

Fact: Intersection of an arbitrary collection of convex sets is convex.

Fact: Intersection of an arbitrary collection of convex sets is convex.

Externally: Convex Hulls:

Let $F \subseteq \mathbb{R}^n$. The intersection of all convex sets containing F is called the convex hull of F and is denoted by $\text{conv}(F)$.

Internally: Convex combinations of $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$:

$$\sum_{i=1}^{k} \lambda_i x^{(i)},$$

for some $\lambda \in \mathbb{R}_+^k$ such that $\sum_{i=1}^{k} \lambda_i = 1$.

Extreme points of convex sets: Let $F \subseteq \mathbb{R}^n$ be a convex set. $x \in F$ is called an extreme point of F if $\nexists u, v \in F \setminus \{x\}$ such that $x = \frac{1}{2}u + \frac{1}{2}v$.

Theorem 1.5 *(Carathéodory's Theorem) Let* $S \subseteq \mathbb{R}^n$ *and* $x \in \text{conv}(S)$*. Then to express* x *as a convex combination of points of* S*,* $(n+1)$ *points of* S *suffice.*

Simplex Method is a family of algorithms based on the above concepts, structures and ideas that solves Linear Optimization problems. It starts from an extreme point of the polyhedron, going from one extreme point to an adjacent one keeps improving the objective function value. It can be proven to run in finite time and it is very efficient in practice; however,

Open Problem: Does any variant of Simplex Method run in polynomial time in $\max\{m, n\}$?

Open Problem: Does any variant of Simplex Method run in polynomial time in $\max\{m, n\}$?

Or even in

$$\max\left\{m, n, \left|\log\left(\max\left\{\left|A_{ij}\right|, \left|b_i\right|, \left|c_j\right|\right\}\right)\right|\right\}?$$

A related open problem involves the graphs constructed from polyhedra. For each extreme point of your polyhedron, make a node in your graph G. Two nodes in G will be connected by an edge if the corresponding extreme points define an edge of the polyhedron.

A related open problem involves the graphs constructed from polyhedra. For each extreme point of your polyhedron, make a node in your graph G. Two nodes in G will be connected by an edge if the corresponding extreme points define an edge of the polyhedron.

Open Problem: Is the length of the longest shortest path between any pair of nodes in G bounded by a polynomial function of $\max\{m, n\}$?

$$
\begin{array}{lllll}
(LP) & \text{Max} & c^T x & & \\
 & & Ax & = & b, \\
 & & x & \geq & 0.
\end{array}
$$

$x \geq 0$ or $x \in \mathbb{R}^n_+$ (x must lie in the nonnegative orthant—a convex cone).

$K \subseteq \mathbb{R}^n$ is a convex cone if $\forall x, v \in K$ and $\alpha \in \mathbb{R}_{++}$, we have $\alpha x \in K$ and $(x + v) \in K$.

Replace $\mathbb{R}^n_+$ by an arbitrary convex cone, to get a more general convex optimization problem.

2 Semidefinite Optimization

Let $\mathbb{S}^n$ denote the space of n-by-n symmetric matrices with entries in $\mathbb{R}$.

Definition 2.1 *Let* $X \in \mathbb{S}^n$.

X *is* positive semidefinite *if*

$$h^T X h \geq 0, \quad \forall h \in \mathbb{R}^n.$$

X *is* positive definite *if*

$$h^T X h > 0, \quad \forall h \in \mathbb{R}^n \setminus \{0\}.$$

For $A, B \in \mathbb{S}^n$, we use the trace inner-product:

$$\langle A, B \rangle := \mathrm{Tr}(A^T B),$$

we write

$$A \succeq B$$

to mean $(A - B)$ is positive semidefinite;

$$A \succ B$$

to mean $(A - B)$ is positive definite.

Note that for $X \in \mathbb{S}^n$ all eigenvalues $\lambda_j(X)$ are real. Also,

$$X \succeq 0 \iff \lambda(X) \geq 0$$

and

$$X \succ 0 \iff \lambda(X) > 0.$$

Given $C, A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, $b \in \mathbb{R}^m$ we have

$$
\begin{array}{lllll}
(P) & \text{inf} & \langle C, X\rangle & & \\
 & & \langle A_i, X\rangle & = & b_i, \quad \forall i \in \{1, 2, \ldots, m\} \\
 & & X & \succeq & 0,
\end{array}
$$

$$
\begin{array}{llll}
(D) & \text{sup} & b^T y & \\
 & & \sum_{i=1}^{m} y_i A_i & \preceq \; C.
\end{array}
$$

We can solve such semidefinite optimization problems

efficiently

both in terms of

- computational complexity theory and
- practical computation.

However, in terms of both computational complexity theory and practical computation the situtation in SDP is much worse than that of LP. (That is, there are very many deep, open problems.)

Open Problem: Which convex optimization problems can be expressed as Semidefinite Optimization problems?

Open Problem: Which convex optimization problems can be expressed as Semidefinite Optimization problems?

Convex sets expressed by multivariate polynomial inequality systems where each polynomial has only real zeros?

3 Separation Theorems

Techniques for proving the Strong Duality Theorem for SDP?

Some elementary topology and real analysis ... and then ... the separation theorem:

$G \subseteq \mathbb{R}^n$ is closed if

$$G = \text{cl}(G) := \left\{ \lim_{k \to +\infty} x^{(k)} : \text{ for some } \left\{ x^{(k)} \right\} \subseteq G \right\}.$$

$G \subset \mathbb{R}^n$ is bounded if

$$G \subseteq \{x \in \mathbb{R}^n : \|x\|_2 \leq R\},$$

for some $R > 0$.

$G \subset \mathbb{R}^n$ is compact if G is closed and bounded.

Theorem 3.1 *(Weierstrass) Let $G \subset \mathbb{R}^n$ be a nonempty compact set and $f : G \to \mathbb{R}$ be a continuous function on G. Then f attains its minimum and maximum values on G.*

Theorem 3.2 *(Separating Hyperplane Theorem) Let* $G \subset \mathbb{R}^n$ *be a nonempty, closed convex set. Suppose* $0 \notin G$*. Then there exist* $a \in \mathbb{R}^n \backslash \{0\}$ *and* $\alpha \in \mathbb{R}_{++}$ *such that*

$$G \subseteq \left\{ x \in \mathbb{R}^n : a^T x \geq \alpha \right\}.$$

In fact, we may take the separating hyperplane to be a supporting hyperplane of G.

Proof: Let $\bar{x} \in G$. Define

$$\bar{G} := \{x \in G : \|x\|_2 \leq \|\bar{x}\|_2\}.$$

Let $a \in \bar{G}$ be the point in $\bar{G}$ with the minimum norm. (Such a point $a \in \mathbb{R}^d$ exists, since $\bar{G} \neq \emptyset$, $\bar{G}$ is compact, $\|x\|_2^2$ is continuous on $\bar{G}$; in fact a is unique since $\|x\|_2^2$ is strictly convex on $\bar{G}$.) Define $\alpha := a^T a > 0$.

For every $x \in G$, the line segment $[a, x]$ lies completely in G. Thus, for every $x \in G$ and for every $\lambda \in (0, 1]$,

$$\begin{aligned} [\lambda x + (1-\lambda)a] \in G \quad &\Rightarrow \quad \|\lambda x + (1-\lambda)a\|_2^2 \geq \|a\|_2^2 = \alpha \\ &\Rightarrow \quad \|\lambda(x-a) + a\|_2^2 \geq \alpha \\ &\Rightarrow \quad \lambda^2\|x-a\|_2^2 + 2\lambda a^T(x-a) \geq 0 \\ &\Rightarrow \quad a^T(x-a) \geq -\frac{\lambda}{2}\|x-a\|_2^2. \end{aligned}$$

Now, we take limits as $\lambda \to 0^+$ and conclude $a^T(x-a) \geq 0$. Therefore, for every $x \in G$, we have $a^T x \geq \alpha$. ■

Theorem 3.3 *(Separating Hypersphere Theorem) Let $G \subset \mathbb{R}^n$ be a nonempty, convex compact set and $\hat{x} \in \mathbb{R}^n \setminus G$. Then there exists a hypersphere which separates $\hat{x}$ from G as follows:*

$G \subseteq \left\{ x \in \mathbb{R}^n : \langle x - \bar{x}, x - \bar{x} \rangle \leq R^2 \right\}$ *and* $\langle \hat{x} - \bar{x}, \hat{x} - \bar{x} \rangle > R^2$.

$\{\alpha v : \alpha \in \mathbb{R}_+\}$ for $v \in K \setminus \{0\}$ defines a ray inside K. Such a ray $R \subseteq K$ is called an extreme ray of K if for every pair of rays $R_1, R_2 \subseteq K$, such that $R_1 + R_2 \supseteq R$ implies either $R_1 = R$ or $R_2 = R$ possibly both. The union of all extreme rays of K is denoted by Ext(K). We also use ext(K) to denote the set of normalized extreme rays of K (where each ray is represented by a single nonzero element of K). For compact, convex sets, we use the notation ext$(\cdot)$ to denote the set of extreme points of the compact, convex set.

A convex set is called pointed if it does not contain any lines.

4 Carathéodory's Theorem revisited for convex cones:

Theorem 4.1 *Let $K \subseteq \mathbb{R}^n$ be a pointed closed convex cone. Then for every $\bar{x} \in K$, there exist $u^{(1)}, u^{(2)}, \ldots, u^{(n)} \in \text{ext}(K)$ and $\lambda \in \mathbb{R}^n_+$ such that*

$$\bar{x} = \sum_{j=1}^{n} \lambda_j u^{(j)}.$$

Example: Let $K := \mathbb{S}^n_+$. Then even though $\dim(K) = \frac{n(n+1)}{2}$, n extreme rays of $\mathbb{S}^n_+$ suffice (Schur/eigenvalue decomposition of $\bar{x}$).

Theorem 4.2 *Let $X \in \mathbb{S}^n$. Then there exists an orthonormal basis $q^{(1)}, q^{(2)}, \ldots, q^{(n)}$ for $\mathbb{R}^n$ such that*

$$X = \sum_{j=1}^{n} \lambda_j q^{(j)} \left(q^{(j)}\right)^T,$$

where λ_j are the eigenvalues of X.

5 Generalizations...

How do we “compute” the convex hull?

First, let's choose a simple but general enough algebraic representation:

Lemma 5.1 *Every compact set $F \subset \mathbb{R}^n$ admits a representation as the feasible region of a system of quadratic inequalities.*

For this lemma, assuming F is closed is enough.

Proof: F is closed. So, $\mathbb{R}^n \setminus F$ is open in $\mathbb{R}^n$; thus, it can be written as the union of (possibly uncountably many) open balls

$$\bigcup_{\bar{x},R} \{x \in \mathbb{R}^n : \|x - \bar{x}\|_2^2 < R^2\} = \mathbb{R}^n \setminus F.$$

By taking the complement of both sides, we express F as the solution set of a system of quadratic inequalities.

$$F = \bigcap_{\bar{x},R} \{x \in \mathbb{R}^n : \|x - \bar{x}\|_2^2 \geq R^2\}.$$

■

Theorem 5.1 *(Kojima, T. [2000]) There is an algorithm based on Semidefinite Optimization which generates a sequence of convex compact sets G_k satisfying*

(a) for every $k \geq 0$, $\text{conv}(F) \subseteq G_{k+1} \subseteq G_k$ (monotonicity), in fact $G_{k+1} = G_k$ iff $G_k = \text{conv}(F)$;

(b) $\bigcap_{k=1}^{k^*} G_k = \emptyset$ for some finite number k^* if $F = \emptyset$ (detecting infeasibility in finitely many steps);

(c) $\bigcap_{k=1}^{\infty} G_k = \text{conv}(F)$ (asymptotic convergence).

Proof: (Proof of (c))

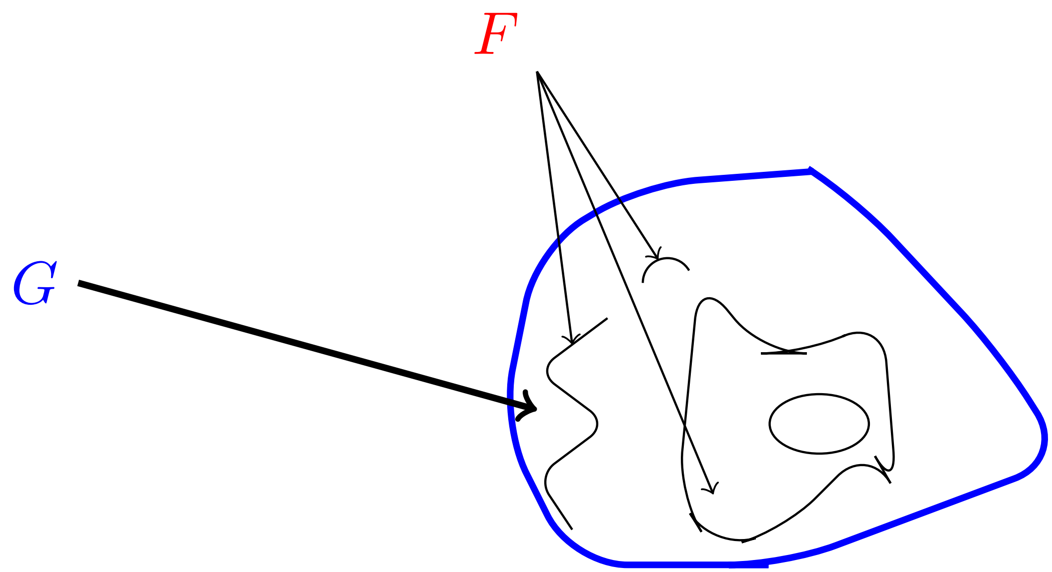

Figure 1: Suppose $G_k \to G \neq \text{conv}(F)$ (seeking a contradiction)

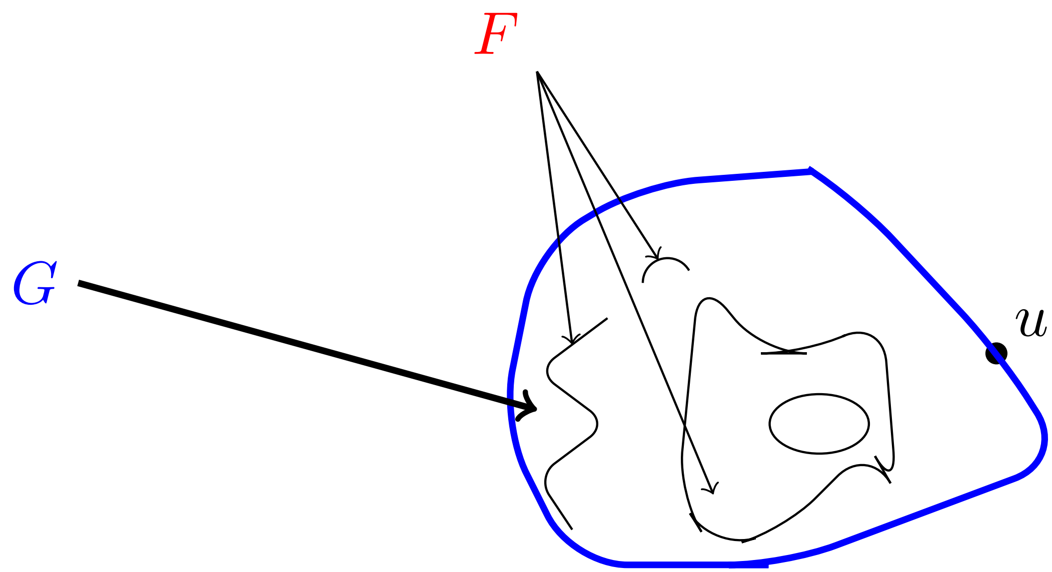

Figure 2: Then, $\exists u \in G \setminus$ conv(F)

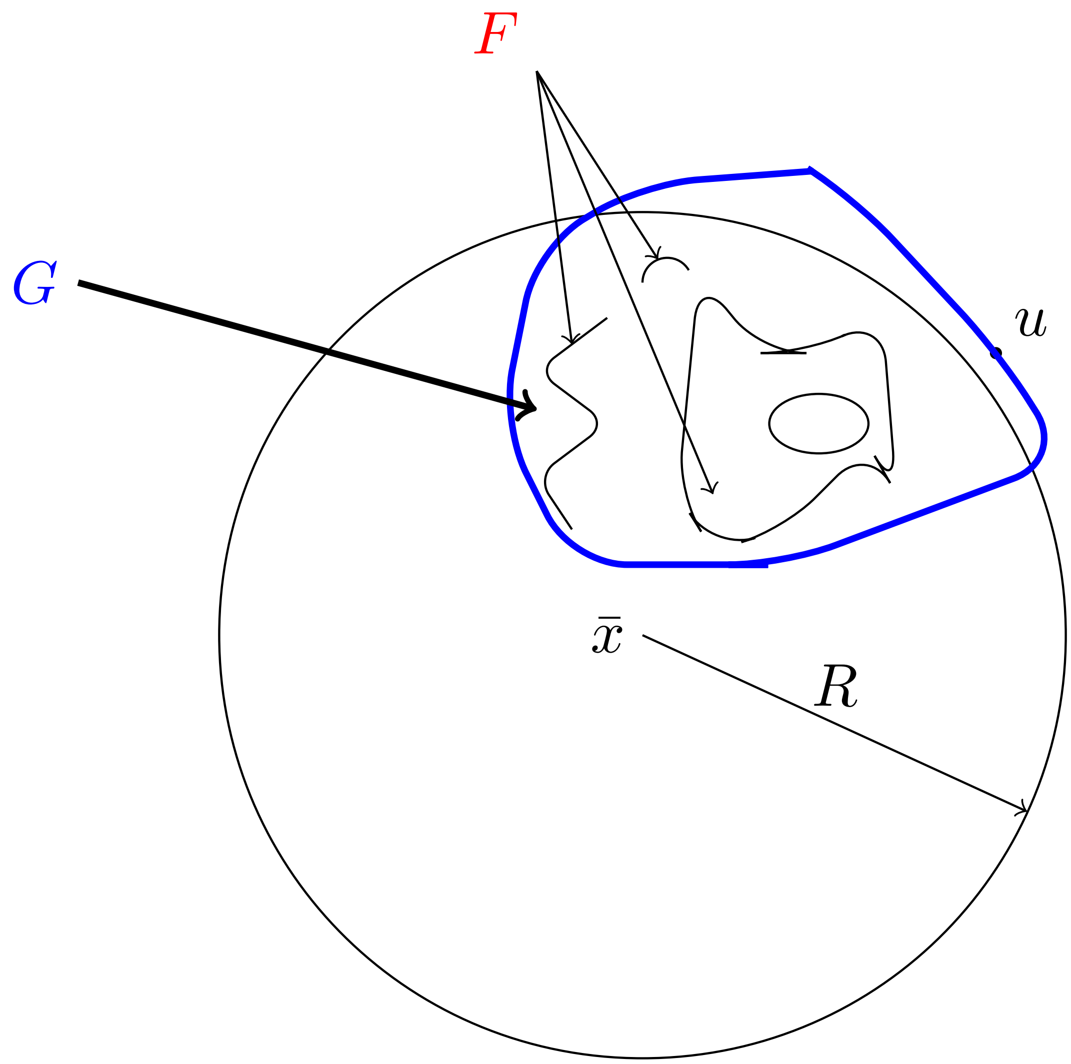

Figure 3: There exists a separating hypersphere for u, conv(F)

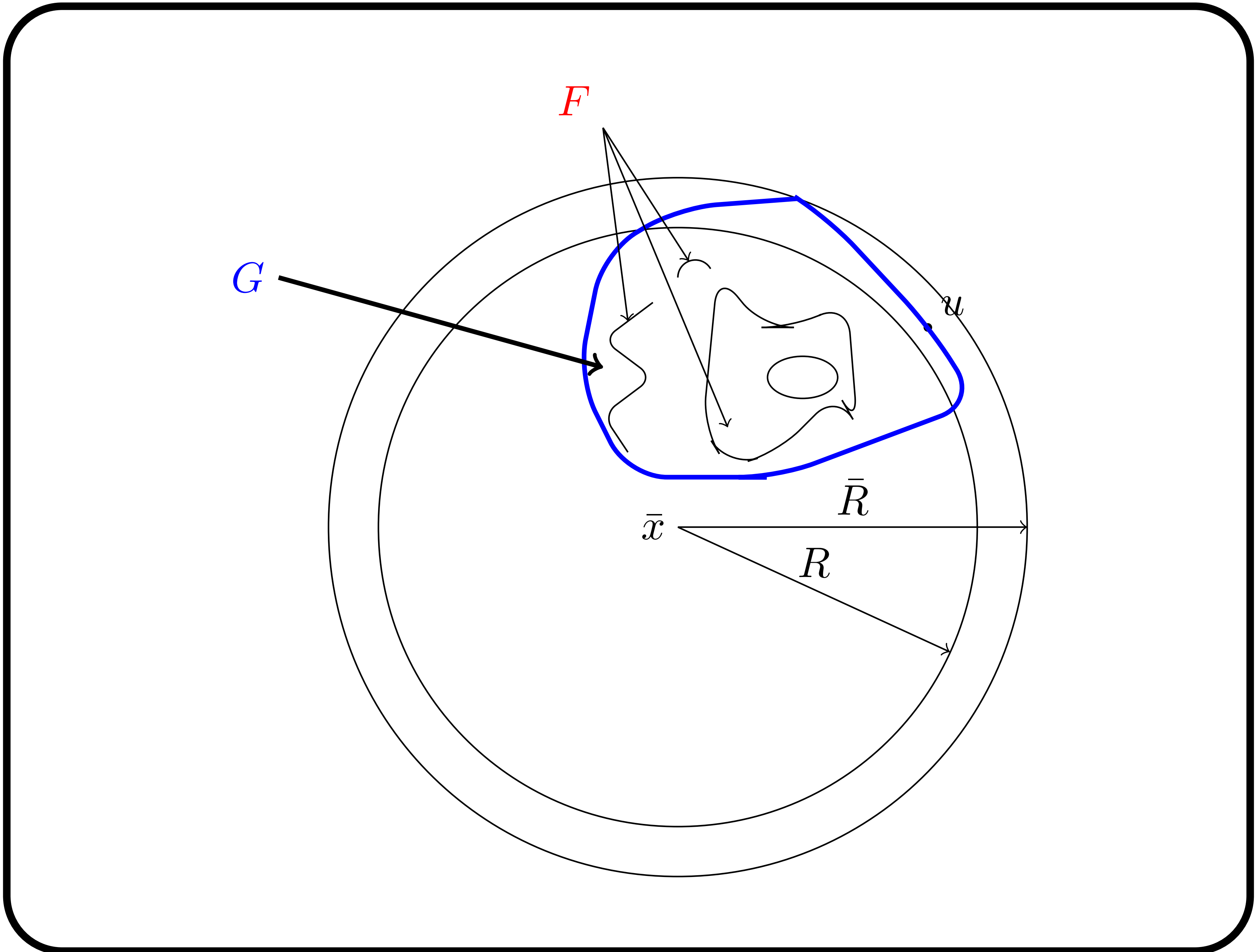

Figure 4: Blow-up the hypersphere to just enclose G

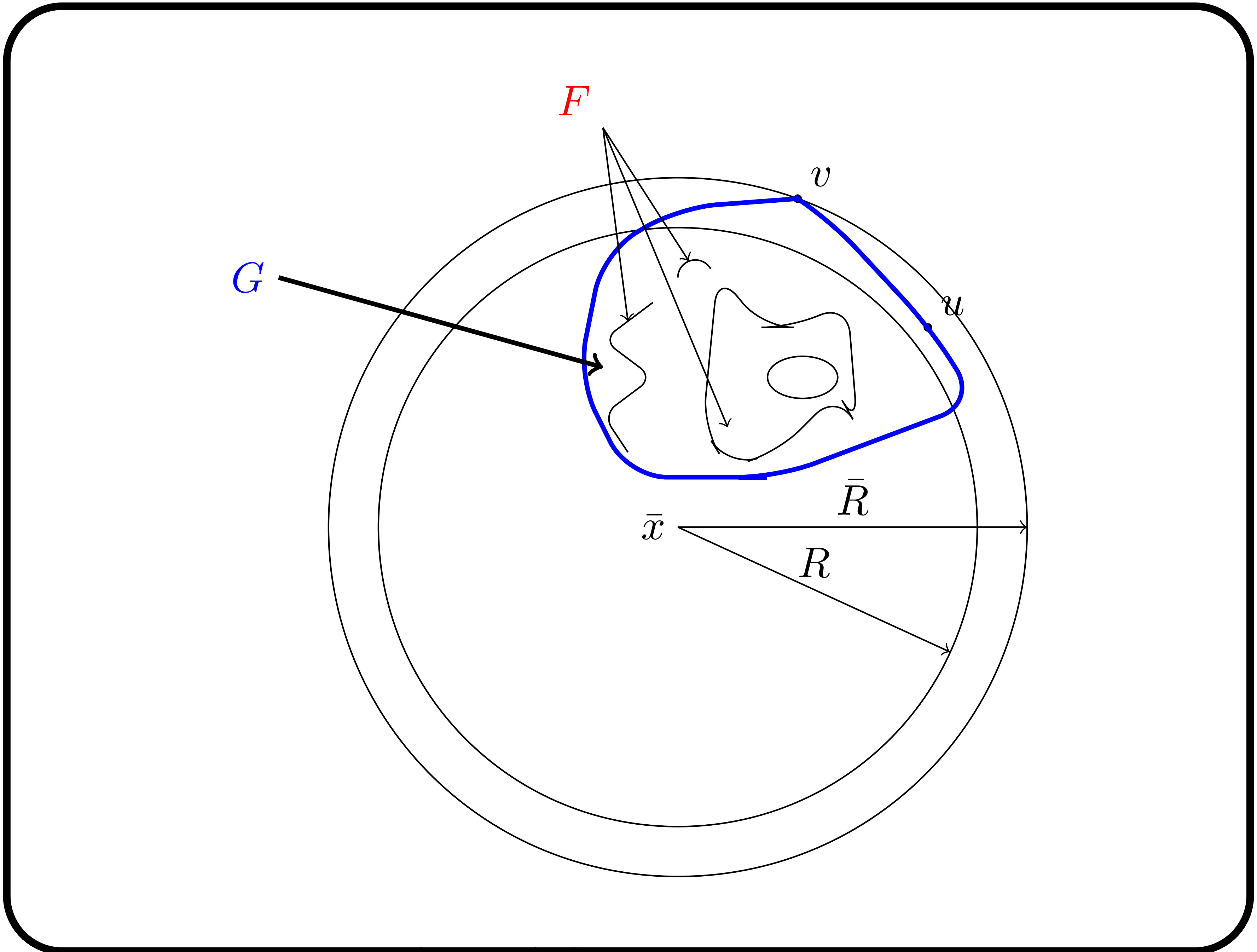

Figure 5: $\exists v \in G \setminus$ conv(F) lying on the blown-up hypersphere

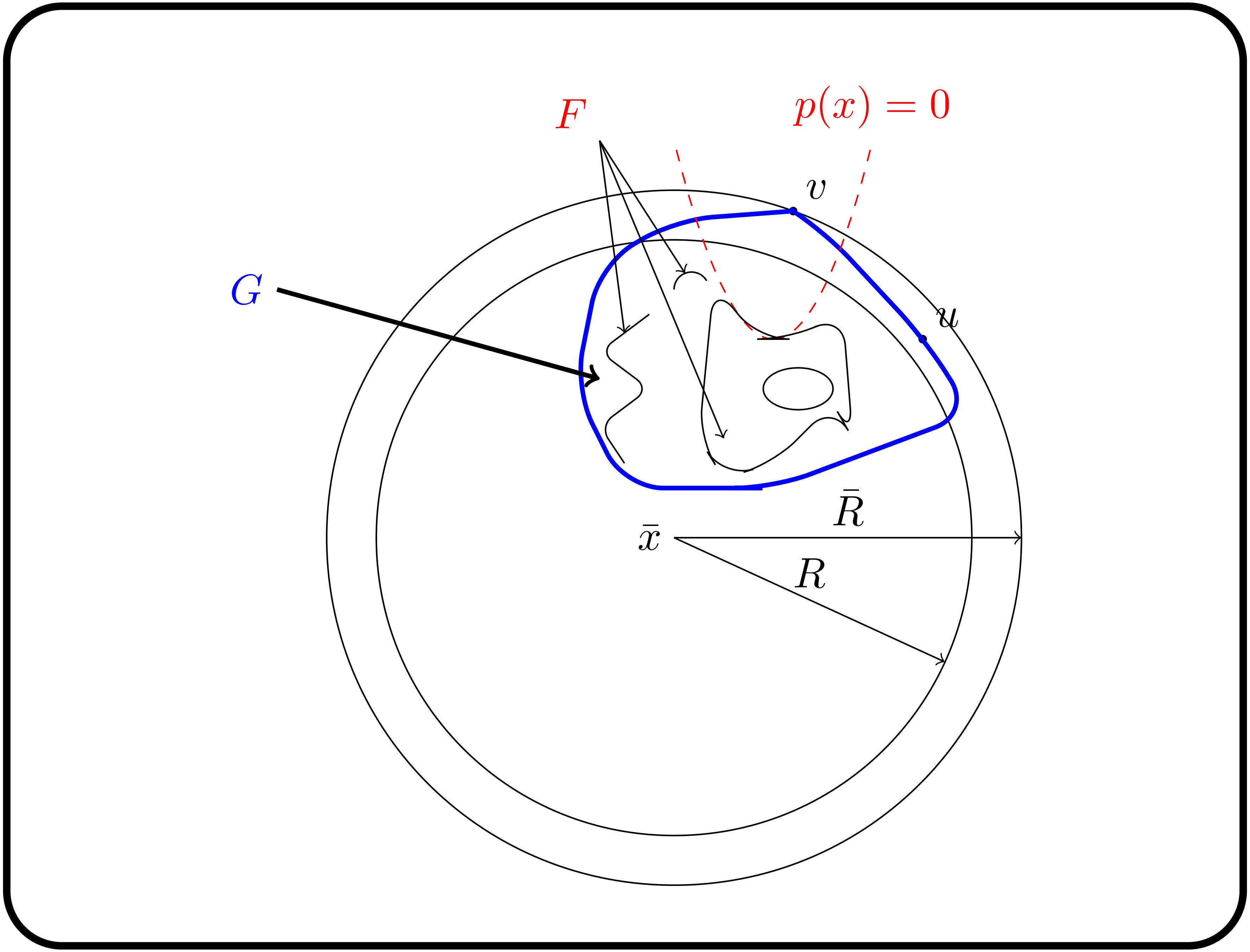

Figure 6: $\exists$ original constraint $p(x) \geq 0$ violated by v

Just a question... How can I convince you that

$$
\begin{aligned}
f(x) \quad := \quad & 83 - 108x_1 + 216x_2 + x_3^2 x_1^2 - 2x_3^3 x_1 - 32x_1 x_2^3 \\
& +24x_1^2 x_2^2 - 8x_1^3 x_2 - 144x_1 x_2^2 + 72x_1^2 x_2 \\
& -216x_1 x_2 + x_3^2 + 54x_1^2 + 216x_2^2 + 432x_1^5 x_2^2 x_4 \\
& -432x_1^4 x_2^3 x_4 + 144x_1^3 x_2^4 x_4 - 576x_1^3 x_2^3 x_4^2 \\
& +256x_1^2 x_2^2 x_4^4 + 864x_1^4 x_2^2 x_4^2 + 768x_1^3 x_2^2 x_4^3 \\
& -16x_1^2 x_2^5 x_4 - 12x_1^3 + 96x_2^3 + x_1^4 + 16x_2^4 + x_3^4 \\
& +96x_1^2 x_2^4 x_4^2 - 256x_1^2 x_2^3 x_4^3 - 12x_1^3 x_2^5 \\
& +54x_1^4 x_2^4 - 108x_1^5 x_2^3 + 81x_1^6 x_2^2 + x_1^2 x_2^6 \\
& +2x_3^2 x_1 - 2x_3^3 \qquad \geq 2, \quad \forall x \in \mathbb{R}^4?
\end{aligned}
$$

What if I claim ...

$$\begin{aligned} f(x) &= (x_1 - 2\,x_2 - 3)^4 + x_3^2\,(x_3 - x_1 - 1)^2 \\ &\quad + x_1^2 x_2^2\,(3\,x_1 - x_2 + 4\,x_4)^4 + 2 \\ &\geq 2, \quad \forall x \in \mathbb{R}^4? \end{aligned}$$

That is, $f(x) =$ Sum-of-Squares $+ 2$, $\forall x \in \mathbb{R}^4$.

6 Applications to Systems of Polynomial Inequalities

Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, e.g., consider the system

$$\begin{aligned} x_1^4 x_2^2 + x_2^3 x_3 + x_1^5 - 1 &\geq 0, \\ 2x_1^3 - x_2^4 &\geq 0 \end{aligned}$$

$$
\begin{aligned}
x_1^4 x_2^2 + x_2^3 x_3 + x_1^5 - 1 &\geq 0, \\
2x_1^3 - x_2^4 &\geq 0
\end{aligned}
$$

which is equivalent to the quadratic system:

$$
\begin{aligned}
x_5 x_6 + x_6 x_7 + x_1 x_5 - 1 &\geq 0, \\
2x_1 x_4 - x_6^2 &\geq 0, \\
x_4 &= x_1^2, \\
x_5 &= x_4^2, \\
x_6 &= x_2^2, \\
x_7 &= x_2 x_3,
\end{aligned}
$$

...

Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, the above results can be translated to the setting of polynomial optimization problems (POP):

$$\begin{array}{lrcl} \min & p_0(x) & & \\ & p_i(x) & \geq & 0, \quad i \in \{1, 2, \ldots, m\}, \end{array}$$

where $p_0, p_1, \ldots, p_m : \mathbb{R}^n \to \mathbb{R}$ are polynomials.

A theorem of Putinar [1993]:

Theorem 6.1 *Suppose* $F := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\}$ *is compact, the polynomials* p_i *have even degree, and their highest degree homogeneous parts do not have common zeroes in* $\mathbb{R}^n$ *except 0. Then every polynomial that is positive on* F *can be written as a nonnegative combination of polynomials of the form*

$$[h_0(x)]^2 + \sum_{i=1}^{m} [h_i(x)]^2 \, p_i(x).$$

Perhaps one of the most fundamental problems here is the *K-moment problem* which is, given $K \subset \mathbb{R}^n$ to decide when a real valued function f of set of monomials in n variables is a moment function $\int_K x^m d\mu$ for some nonnegative Borel measure μ on K. Schmüdgen [1991] characterized the solutions to the K-moment problem (called *K-moment sequences*) for all compact semi-algebraic sets K in terms of the positive definiteness of matrices arising from the moment functions. Schmüdgen's proof utilizes Positivstellensatz in proving the above-mentioned algebraic fact. The result also generalizes many other preexisting beautiful results such as Handelman's Theorem [1988]; some of these connections are old and they generalize results some of which go all the way back to Minkowski in late 1800's.

Positivstellensatz (Stengle [1974]):

$F := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\} = \emptyset$ iff that there exists $g \in$ cone $(p_1, p_2, \ldots, p_m)$ such that $g(x) = -1$.

That is, iff there exist $s_0, s_J, \ldots \in$ SoS$(n, *)$ such that

$$g = \sum_{J \subseteq \{1,2,\ldots,m\}} s_J \prod_{i \in J} p_i = -1.$$

Positivstellensatz is a "common" generalization of Farkas' Lemma

Lemma 6.1 *(Farkas' Lemma) Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ be given. Then exactly one of the following systems has a solution:*

(I) $Ax = b, x \geq 0$;

(II) $A^T y \geq 0, b^T y < 0.$

and Hilbert's Nullstellensatz [1901] (characterizing when a system of polynomial equations has no solution over $\mathbb{C}^n$): Only for this slide, let p_i be polynomials in complex variables ($x \in \mathbb{C}^n$). Then exactly one of the following systems has a solution:

(I) $p_i(x) = 0, \forall i \in \{1, 2, \ldots, m\}$;

(II) $\exists$ polynomials h_i such that $\sum_{i=1}^m h_i(x) p_i(x) = -1$.

Positivstellensatz leads to SoS certificates using Convex or Semidefinite Optimization!

Given $x \in \mathbb{R}^n$ and polynomial $f : \mathbb{R}^n \to \mathbb{R}$ of degree $2d$, let

$$h(x) := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, x_1 x_3, \ldots, x_2^2, \ldots, x_n^d]^T \in \mathbb{R}^N,$$

where $N := \binom{n+d}{d}$. We are interested in

$$\mathcal{F}(f) := \left\{ X \in \mathbb{S}^N : [h(x)]^T X h(x) = f(x) \right\}.$$

The following well-known fact connects SoS and semidefinite optimization.

Theorem 6.2 *Let $\bar{z} \in \mathbb{R}$. Then $[f(x) - \bar{z}]$ is SoS iff*

$$\{X \in \mathcal{F}(f) : X \succeq \bar{z} e_1 e_1^T\} \neq \emptyset.$$

7 Final Remarks

- We took a tour in mathematical optimization focusing on some aspects that are related to some of my research interests. The area is much, much wider...

- We took a tour in mathematical optimization focusing on some aspects that are related to some of my research interests. The area is much, much wider...
- I wanted to give you some idea of some of the techniques and connections.

- We took a tour in mathematical optimization focusing on some aspects that are related to some of my research interests. The area is much, much wider...
- I wanted to give you some idea of some of the techniques and connections.
- Beautiful, interesting, useful with deep connections to other areas of mathematics and in general mathematical sciences...

- We took a tour in mathematical optimization focusing on some aspects that are related to some of my research interests. The area is much, much wider...
- I wanted to give you some idea of some of the techniques and connections.
- Beautiful, interesting, useful with deep connections to other areas of mathematics and in general mathematical sciences...
- Also, there are very many interesting open problems...