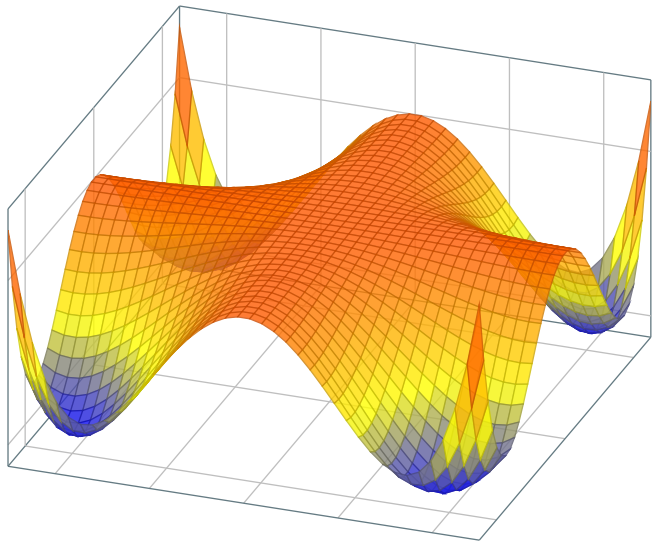
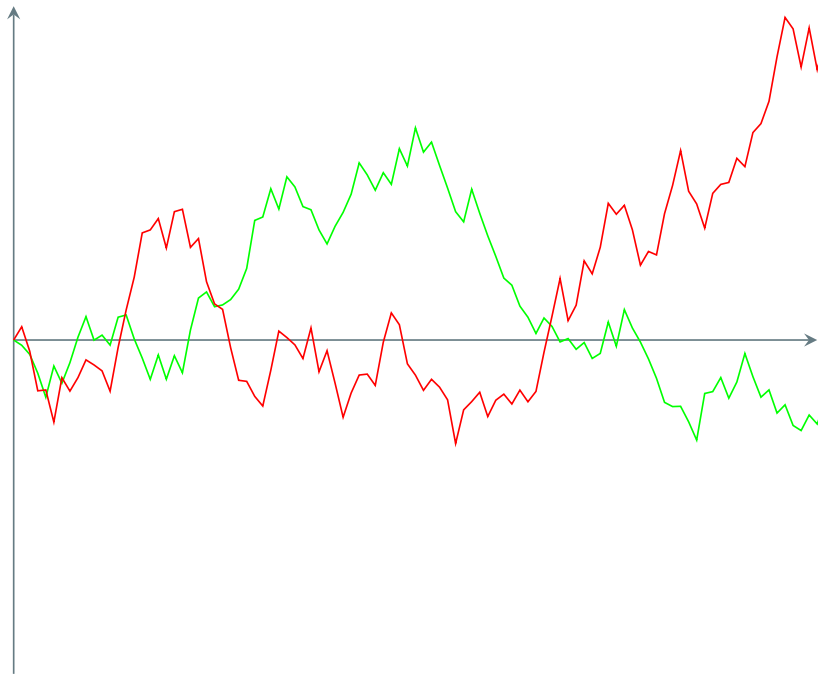
Real algebra, random walks, and information theory

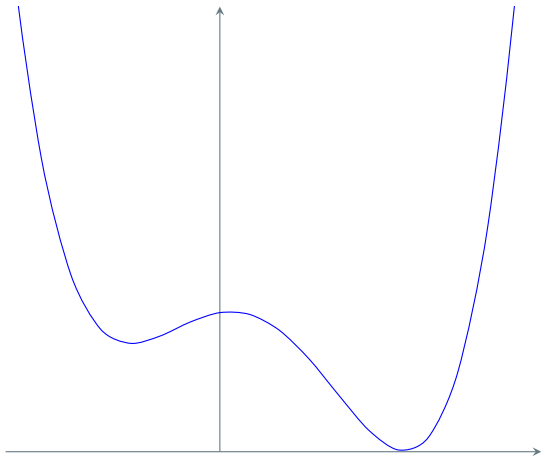
Tobias Fritz

March 2019





- When does a polynomial in $f \in \mathbb{R}[X]$ take on only nonnegative values?

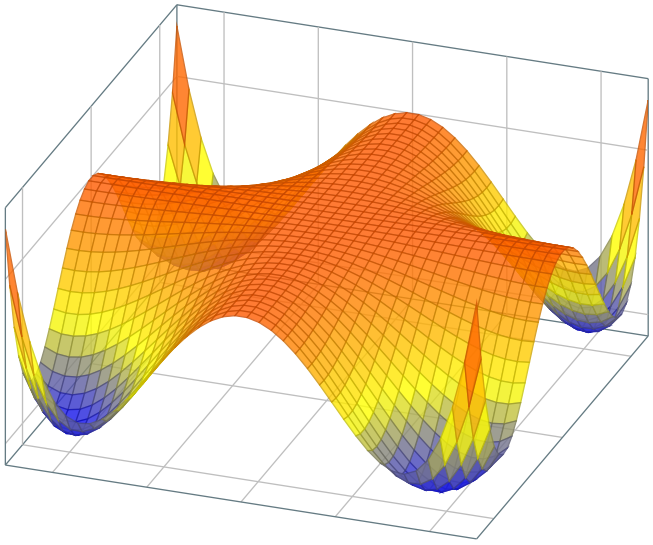


$$f = X^4 - 2X^3 - 6X^2 + 2X + 25$$
$$= (X^2 - X - 4)^2 + (X - 3)^2$$

- If and only if f is a **sum of squares** of polynomials.
  Proof by fundamental theorem of algebra!

- Writing f as a sum of squares is a **certificate** of nonnegativity.

- When is $f \in \mathbb{R}[X, Y]$ nonnegative? Example: **Motzkin polynomial**



$$M := X^4Y^2 + X^2Y^4 + 1 - 3X^2Y^2$$

$$= 3\left(\frac{X^4Y^2 + X^2Y^4 + 1}{3} - \sqrt[3]{(X^4Y^2)\cdot(X^2Y^4)\cdot 1}\right)$$

$$\geq 0.$$

- *M cannot* be written as a sum of squares of polynomials.

- *M can* be written as a sum of squares of **rational** functions. Also a certificate of nonnegativity!

Hilbert's 17th problem

Theorem (Artin '27)

Every multivariate polynomial $f \in \mathbb{R}[\underline{X}]$ with $f \geq 0$ can be written as a sum of squares of rational functions:

$$f = \frac{g_1^2 + \ldots + g_m^2}{h^2}$$

for $g_1, \ldots, g_m, h \in \mathbb{R}[\underline{X}]$.

- ▷ Surprisingly, no known proof without model theory!

Real algebra

- More generally, real algebra studies the relation between:
  - geometric positivity = taking nonnegative values,
  - algebraic positivity = existence of a nonnegativity certificate.

  A **Positivstellensatz** gives conditions for when the two coincide. For $\mathbb{R}[X]$ or for classes of **ordered rings**.

- Many results are about nonnegativity instead of positivity (sometimes *Nichtnegativstellensatz*).

- Scary terminology due to analogy with Hilbert's *Nullstellensatz*.

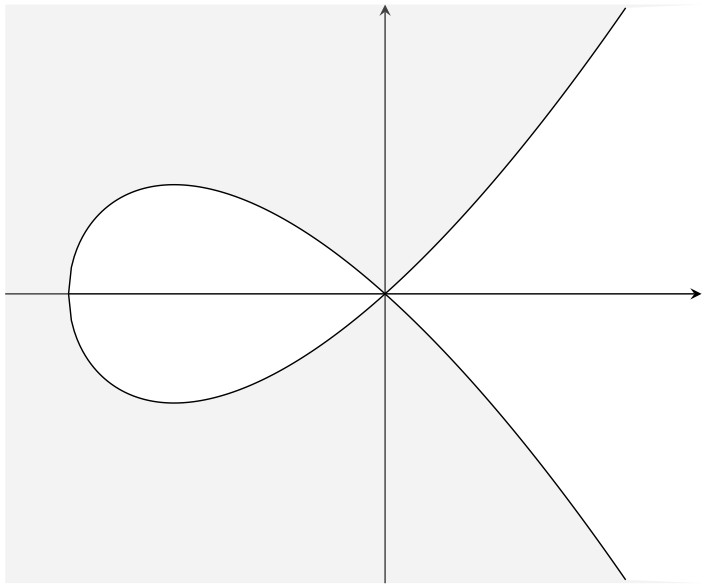
- One Nullstellensatz, many Positivstellensätze!

Example: Stengle's Positivstellensatz

- How about results for positivity of polynomials on subsets $S \subseteq \mathbb{R}^n$?
- So let $\{g_1, \ldots, g_n\}$ be a finite set of polynomials in $\mathbb{R}[\underline{X}]$, and

$$S := \{x \in \mathbb{R}^d \mid g_i(x) \geq 0\}.$$

- Running example:



$$S = \{(x, y) \in \mathbb{R}^2 \mid y^2 - x^2(x+1) \geq 0\}$$

▷ Try to certify nonnegativity of $f \in \mathbb{R}[\underline{X}]$ by proving membership in

$$P(g_1, \ldots, g_n) := \mathrm{lin}_{\mathbb{R}_+} \left\{ h^2 g_1^{\beta_1} \cdots g_m^{\beta_m} \ : \ h \in \mathbb{R}[\underline{X}],\ \beta_i \in \{0,1\} \right\}.$$

Stengle's Positivstellensatz (Krivine '64, Stengle '74)

$$f > 0 \text{ on } S \quad \Longleftrightarrow \quad \exists p, q \in P(g_1, \ldots, g_n),\ f = \frac{1+p}{q}.$$

▷ Example: $g_1 = Y^2 - X^2(X+1)$ from before, and

$$f = Y^2 - X + 1.$$

Then $f \notin P(g_1)$, but

$$f = \frac{1 + \overbrace{(Y^2 - X^2(X+1))}^{g_1} + \overbrace{(X^2Y^2 + X^2)}^{\text{sum of squares}}}{\underbrace{X^2+1}_{\text{sum of squares}}}$$

Motivating application:

- ▷ $\mathrm{Rep}(G)$, the **semiring** of finite-dimensional complex representations of a compact Lie group G, ordered by inclusion of subrepresentations.

- ▷ Recall: a semiring is like a ring, but need not have additive inverses. E.g.: $\mathbb{N}$, tropical semiring, ...

- ▷ Instead of positivity, now consider *comparison* $x \leq y$ of semiring elements x and y, both geometrically and algebraically.

- ▷ Motivating question: for given U and V, when is there $n \in \mathbb{N}_{>0}$ with

$$U^{\otimes n} \hookrightarrow V^{\otimes n} \quad ?$$

  $\Rightarrow$ An algebraic comparison.

- Focus on $G = SU(2)$ for simplicity.

- $\mathrm{Rep}(SU(2))$ is isomorphic to the polynomial semiring $\mathbb{N}[X]$ equipped with the "twisted" multiplication

$$X^n \cdot X^m := \sum_{j=|n-m|}^{n+m\,\prime} X^j \qquad \text{(only every second term in sum)},$$

  with the coefficientwise ordering.

- Example: We have

$$\begin{aligned} X &\not\leq X^3 \\ X \cdot X &\leq X^3 \cdot X^3 \\ [\text{since} \quad 1 + X^2 &\leq 1 + X^2 + X^4 + X^6\,] \end{aligned}$$

  $\Rightarrow$ The question about $U^{\otimes n} \hookrightarrow V^{\otimes n}$ is nontrivial.

- ▷ Existing Positivstellensätze *do not apply*.
- ▷ I will now state a Positivstellensatz which does, but answers a slightly modified question.
- ▷ It generalizes Strassen's Positivstellensatz[1].

Definition

An element u in a semiring S is **polynomially universal** if for every nonzero $x \in S$ there is $p \in \mathbb{N}[X]$ such that

$$x \leq p(u), \qquad 1 \leq xp(u).$$

[1] Volker Strassen. "The asymptotic spectrum of tensors". In: **J. Reine Angew. Math.** 384 (1988), pp. 102–152.

Theorem (T.F. '18[2], slightly simplified)

Let S be an ordered semiring with $\mathbb{Q}_+ \subseteq S$ and polynomially universal $u \in S$. Then for nonzero $x, y \in S$, the following are equivalent:

(a) $f(x) \leq f(y)$ for all monotone homomorphisms $f : S \to \mathbb{R}_+$;

(b) For all $r \in \mathbb{R}_+$ and $\varepsilon > 0$, there is $p \in \mathbb{Q}_+[X]$ with $p(r) \leq 1 + \varepsilon$ and nonzero $z \in S$ such that

$$zx \leq p(u)\, zy.$$

(c) For all $r \in \mathbb{R}_+$ and $\varepsilon > 0$, there are $p \in \mathbb{N}[X]$ and $n \in \mathbb{N}_{>0}$ such that $p(r) \leq (1 + \varepsilon)^n$ and

$$x^n \leq p(u) y^n.$$

---

[2] Tobias Fritz. **A generalization of Strassen's Positivstellensatz and its application to large deviation theory**. arXiv:1810.0866.

The proof uses:

- A reduction to the semifield case;
- A functional analysis argument showing that the cone of monotone functionals $S \to \mathbb{R}$ is spanned by its extreme rays;
- Some tricks to show that these extreme rays are exactly the homomorphisms in the semifield case.

Semifields themselves are interesting:

Example

For a compact Hausdorff space X,

$$S := C(X, \mathbb{R}_{>0}) \cup \{0\}$$

is a semifield which does not embed into a field.

- For the application, need to understand the monotone homomorphisms $\mathrm{Rep}(SU(2)) \to \mathbb{R}_+$.

Theorem (Székelyhidi[3])

The monotone homomorphisms $f : \mathrm{Rep}(SU(2)) \to \mathbb{R}_+$ are exactly the maps

$$f_\lambda\left(\sum_n a_n X^n\right) := (\sinh\lambda)^{-1}\sum_n a_n \sinh((n+1)\lambda),$$

parametrized by $\lambda \in \mathbb{R}_+$.

[3] László Székelyhidi. **Functional equations on hypergroups**. World Scientific, 2013.

Theorem (T.F., work in progress)

Let U and V be representations of $SU(2)$. Then the following are equivalent:

(a) $f_\lambda(U) \leq f_\lambda(V)$ for all $\lambda \in \mathbb{R}_+$.

(b) For all $\lambda \in \mathbb{R}_+$ and $\varepsilon > 0$ there exist a representation W and $n \in \mathbb{N}_{>0}$ such that $f_\lambda(W) \leq (1+\varepsilon)^n$ and

$$U^{\otimes n} \hookrightarrow V^{\otimes n} \otimes W.$$

Second application

- Let Meas($\mathbb{R}$) be the semiring of compactly supported measures on $\mathbb{R}$ under convolution as multiplication,

$$\int f(x)\, d(\mu * \nu)(x) := \int \int f(x+y)\, d\mu(x)\, d\nu(y).$$

- It is ordered by the **usual stochastic order**, where $\mu \geq \nu$ means

$$\mu([c,\infty)) \geq \nu([c,\infty)) \quad \forall c \in \mathbb{R}.$$

- Measures on $\mathbb{Z}$ form a subsemiring isomorphic to $\mathbb{R}_+[X, X^{-1}]$,

$$\sum_{n\in\mathbb{Z}} a_n X^n \quad \longmapsto \quad \sum_{n\in\mathbb{Z}} a_n \delta_n,$$

with order generated by $X \geq 1$.

Proposition

The monotone homomorphisms $f_t : \mathrm{Meas}(\mathbb{R}) \to \mathbb{R}_+$ are exactly the values of the **moment-generating function**

$$f_t(\mu) = \int e^{tx}\, d\mu(x)$$

parametrized by $t \in \mathbb{R}_+$.

- On $\mathbb{R}[X, X^{-1}]$, this restricts to evaluation at e^t,

$$f_t\left(\sum_{n\in\mathbb{Z}} a_n \delta_n\right) = \int e^{tx}\, d\left(\sum_{n\in\mathbb{Z}} a_n \delta_n\right) = \sum_n a_n (e^t)^n.$$

- Applying our Positivstellensatz gives a comparison theorem for random walks!

Theorem (T.F. '18 [4])

For bounded random variables X and Y, the following are equivalent:

(a) For all $t \in \mathbb{R}_+$,

$$\mathbb{E}[e^{tX}] \geq \mathbb{E}[e^{tY}].$$

(b) For every $r \in \mathbb{R}_+$ and $\varepsilon > 0$, there exists bounded W, making all variables independent, and $n \in \mathbb{N}$ such that $\mathbb{E}[e^{rW}] \leq (1+\varepsilon)^n$

$$\mathbf{P}\left[W + \sum_{i=1}^{n} X_i \geq c\right] \geq (1-\varepsilon)^n \, \mathbf{P}\left[\sum_{i=1}^{n} Y_i \geq c\right] \qquad c \in \mathbb{R}.$$

- Taking Y to be constant recovers a version of Cramér's large deviation theorem!

Submajorization

Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_m)$ be vectors with entries in $\mathbb{R}_+$.

Definition

a **submajorizes** b, denoted $a \succ_w b$, if

$$\sum_{j=1}^{k} a_i \geq \sum_{j=1}^{k} b_i$$

for all k, assuming decreasing rearrangement, $a_1 \geq \ldots \geq a_n$ and $b_1 \geq \ldots \geq b_m$, and padding with 0's if necessary.

Theorem (Ky Fan)

For $A, B \in M_n(\mathbb{C})$, the vectors of singular values satisfy

$$s(A + B) \prec_w s(A) + s(B).$$

- Vectors up to permutation and padding zeroes: ordered semiring Major, with:
  - direct sum $a \oplus b$ as addition;
  - Schur product $a \otimes b$ as multiplication;
  - the submajorization order.

Proposition

The monotone homomorphisms Major $\to \mathbb{R}_+$ are exactly the ℓ^p-norms for $p \in [1, \infty)$,

$$\|a\|_p := \sum_i a_i^p.$$

- Of particular interest is the case of probability vectors a with $\sum_i a_i = 1$. For these, submajorization = **majorization**.

Catalytic and asymptotic submajorization

Theorem (T.F., work in progress)

For vectors a and b, the following are equivalent:

(a) $\|a\|_p \geq \|b\|_p$ for all $p \in [1, \infty)$;

(b) For every $p \in [1, \infty)$ and $\varepsilon > 0$, there are $c, d \in \text{Major}$ such that $\|d\|_p \leq 1 + \varepsilon$, and

$$a \otimes c \otimes d \succ_w b \otimes c.$$

(c) For every $p \in [1, \infty)$ and $\varepsilon > 0$, there are $n \in \mathbb{N}$ and $d \in \text{Major}$ such that $\|d\|_p \leq (1 + \varepsilon)^n$ and

$$a^{\otimes n} \otimes d \succ_w b^{\otimes n}.$$

Very similar to results of Aubrun and Nechita[5], with some differences:

- ▷ minor: normalization of probability,
- ▷ major: the role of d is replaced by the closure in ℓ^1-norm.

Problem: Recover their results?

Remark (essentially Aubrun and Nechita)

The **surprisal map**

$$I : \mathrm{Major} \longrightarrow \mathrm{Meas}(\mathbb{R}), \qquad a \longmapsto \sum_{i\,:\,a_i>0} a_i\,\delta_{\log a_i}$$

is a semiring homomorphism which preserves and reflects the order,

$$P \succ_w Q \quad \Longleftrightarrow \quad I(P) \geq I(Q).$$

[5] Guillaume Aubrun and Ion Nechita. "Catalytic majorization and ℓ_p norms". In: **Comm. Math. Phys.** 278.1 (2008). arXiv:quant-ph/0702153, pp. 133–144.

General remarks

- Most standard applications:

  $$\text{Find algebraic certificate} \implies \text{Conclude geometric nonnegativity}$$

  Example: Polynomial optimization via semidefinite programming[6].

- My applications:

  $$\text{Detect geometric nonnegativity} \implies \text{Algebraic certificate exists}$$

- Traditionally, emphasis on polynomial rings $\mathbb{R}[\underline{X}] = \mathbb{R}[X_1, \ldots, X_d]$.

- My applications: other **ordered rings** and **ordered semirings**.

---

[6] Jean Bernard Lasserre. **An introduction to polynomial and semi-algebraic optimization**. Cambridge University Press, Cambridge, 2015.

The next steps

- For real algebra:
  - Prove a Positivstellensatz for *modules* over an ordered ring A, specializing to the Hahn-Banach separation theorem in case $A = \mathbb{R}$. ✓
  - Achieve some unification of the myriad of Positivstellensätze. (✓)
  - Find a new proof of Stengle's Positivstellensatz avoiding the use of model theory. (Just a dream.)
  - Potential noncommutative version of the Positivstellensatz?

The next steps

- For applications:
  - Get better bounds on the "fudge factors", e.g. by a sharpened Positivstellensatz.
  - Application to characterizations of Laplace transforms. ✓
  - Application to Shannon's mathematical theory of communication. (✓)
  - Ideas for further applications?

The End

Bonus: The mathematical theory of communication

Definition (Shannon '48[7])

(a) An **alphabet** is a finite set.

(b) For alphabets A and B, a **communication channel** $P : A \to B$ is a stochastic matrix $(P_{ab})_{a \in A, b \in B}$,

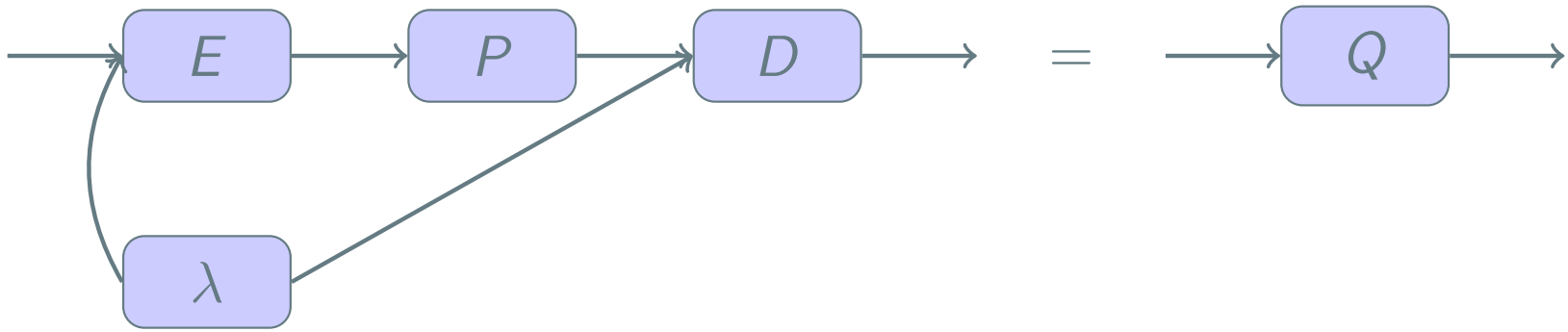
$$P_{ab} \geq 0, \qquad \sum_b P_{ab} = 1.$$

- ▷ P_{ab} = probability that symbol b is received if symbol a is sent.
- ▷ The **perfect channel** is the identity matrix.
- ▷ We are interested in using **codes** to use one channel to simulate another, in particular a perfect channel.

---

[7] Claude E. Shannon. "A mathematical theory of communication". In: **Bell System Tech. J.** 27 (1948), pp. 379–423, 623–656.

- So we put $P \geq Q$ if P can be used to simulate Q in the presence of shared randomness[8],



- E = encoder, D = decoder, λ = shared randomness.
- We have direct sum and tensor product of stochastic matrices.
- Operational significance: nondeterministic choice and parallel use of channels.
- $\Rightarrow$ Ordered semiring CommCh.

[8] Claude E. Shannon. "A note on a partial ordering for communication channels". In: **Information and Control** 1 (1958), pp. 390–397.

- **Problem:** In contrast to previous examples, the monotone homomorphisms $\mathrm{CommCh} \to \mathbb{R}_+$ are very difficult to classify!

- Surprisingly, exponentiating Shannon's **channel capacity**

$$C(P) := \sup_X I(X : P(X))$$

  gives one of them.

- Others arise from graph invariants and combinatorial geometries[9], including the Lovász number and fractional chromatic number.

- $\Rightarrow$ At current state of the art, applying the Positivstellensatz does not tell us much.

---

[9] Tobias Fritz. **A unified construction of semiring-homomorphic graph invariants**. arXiv:1901.01090.

Bonus: The moment problem for modules

- ▷ We also have a Positivstellensatz for **modules**.
- ▷ Let A be an ordered $\mathbb{Q}$-algebra and M an A-module. M is **ordered** if it comes equipped with a positive cone $M_+ \subseteq M$ such that $A_+ M_+ \subseteq M_+$.

Definition

(a) $u \in A$ is **polynomially M-absorbent** if for every $a \in A_+$ there is $p \in \mathbb{Q}_+[X]$ such that

$$a\, m \leq p(u)\, m$$

for every $m \in M_+$.

(b) $v \in M_+$ is an **A-order unit** if for every $m \in M$ there is $a \in A$ such that

$$m \leq a\, v.$$

If both exist, then M is **of polynomial growth**.

Theorem

Let M be of polynomial growth over A and $f : M \to \mathbb{R}$ positive and continuous[10]. Then the following are equivalent:

(a) For every $a \in A_+$ there is a positive $\mathbb{Q}$-linear map $f' : M \to \mathbb{R}_+$ such that for all $m \in M$,

$$f(m) = f'((1 + a)m).$$

(b) For all $a \in A$ and $m \in M_+$,

$$f(a^2\, m) \geq 0.$$

(c) For all $m \in M_+$ and $a_1, \ldots, a_n \in A_+$, the matrix

$$(f(a_i a_j m))_{i,j=1}^{n}$$

is positive semidefinite.

---

[10] In locally convex topology generated by $\mathcal{N}_\varepsilon := \bigcup_n \left[-\sum_{k=0}^{n} \varepsilon^k u^k v, \sum_{k=0}^{n} \varepsilon^k u^k v\right]$.

(d) There is a Hilbert space $\mathcal{H}$ together with a positive $\mathbb{Q}$-linear homomorphism $\pi : A \to \mathcal{B}\mathcal{H}$ and a positive A-linear continuous map $\eta : M \to \pi(A)''$ as well as a vector $\xi \in \mathcal{H}$ such that

$$f(m) = \langle \xi, \eta(m)\, \xi \rangle$$

for all $m \in M$ and $\eta(v) = 1$.

(e) There is a compactly supported Radon measure ν on $\mathrm{Sper}(M)_{\mathbb{R}}$ such that

$$f(m) = \int_{\mathrm{Sper}(M)_{\mathbb{R}}} \psi(m)\, d\nu(\phi, \psi).$$

(f) There is a compactly supported Radon measure μ on $\mathrm{Sper}(A)_{\mathbb{R}}$ and a positive A-module map $\mathbb{E} : M \to L^{\infty}(\mathrm{Sper}(A)_{\mathbb{R}}, \mu)$ such that

$$f(m) = \int \mathbb{E}[m]\, d\mu.$$

for all $m \in M$, and $\mathbb{E}[v] = 1$.

Remarks:

- ▷ This result can be dualized, giving the first Positivstellensatz for modules, specializing to a version of the Hahn–Banach separation theorem.

- ▷ Specializing to $M = A$: gives more specific Positivstellensatz, from which both Putinar's and our earlier one are easy to derive.

- ▷ Open problem: remove the continuity assumption on f.

Application to the Laplace transform:

Definition

A function $f : \mathbb{R}^d \to \mathbb{R}$ is a **Laplace transform** if there is a positive measure μ on $\mathbb{R}_+^n$ such that

$$f(x) = \int e^{\langle t,x \rangle} \, d\mu(t).$$

- For $d = 1$, we get the usual

$$f(x) = \int_0^\infty e^{tx} \, d\mu(t).$$

Theorem (T.F., work in progress)

Let $f : \mathbb{R}^d \to \mathbb{R}$ be nonnegative, monotone, of exponential growth. Then the following are equivalent:

(a) f is a Laplace transform;

(b) f has deconvolutions: for every nonzero $\mu \in \mathrm{Meas}(\mathbb{R}^d)$, there is measurable nonnegative monotone $g : \mathbb{R}^d \to \mathbb{R}$ such that

$$f(x) = \int_{\mathbb{R}^d} g(x+t)\, d\mu(t).$$

(c) For any compactly supported **signed** measure ν on $\mathbb{R}^d$, the function

$$x \longmapsto \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x+y+z)\, d\nu(y)\, d\nu(z)$$

is monotone.

(d) For every $x, y \in \mathbb{R}^d$ with $y - x \in C$ and $z_1, \ldots, z_m \in \mathbb{R}^d$, the matrix

$$\left(f(z_i + z_j + x) - f(z_i + z_j + y)\right)_{i,j=1}^{m}$$

is positive semidefinite.

- ▷ This can be generalized further to arbitrary closed generating cones $C \subseteq \mathbb{R}^d$ instead of $\mathbb{R}^d_+ \subseteq \mathbb{R}^d$.
- ▷ Exponential growth $\simeq$ compact support of the measure.
- ▷ Would like to eliminate this assumption!
- ▷ I do not yet know which of these characterizations is new.