A constructive version of Carathéodory's Convexity Theorem

Josef Berger Gregor Svindland

May 5, 2021

Abstract

Carathéodory's Convexity Theorem states that each element in the convex hull of a subset A of $\mathbb{R}^m$ can be written as the convex combination of $m+1$ elements of A. We prove an approximate constructive version of Carathéodory's Convexity Theorem for totally bounded sets.

For each $n \in \mathbb{N}$, define

$$I_n := \{1, \ldots, n\}.$$

For any $\lambda \in \mathbb{R}^n$, we denote by λ_i the ith coordinate of λ, that is $\lambda = (\lambda_1, \ldots, \lambda_n)$. Let

$$S_n := \left\{ \lambda \in \mathbb{R}^n \mid \forall i \in I_n \, (0 \leq \lambda_i) \wedge \sum_{i \in I_n} \lambda_i = 1 \right\}.$$

The linear space generated by $x^1, \ldots, x^n \in \mathbb{R}^m$ is denoted by

$$\operatorname{span}(\{x^1, \ldots, x^n\}) := \left\{ \sum_{i \in I_n} \lambda_i x^i \mid \lambda \in \mathbb{R}^n \right\}.$$

The *convex hull* of an inhabited subset A of $\mathbb{R}^m$—that is there exists $x \in A$—is

$$\operatorname{co}(A) = \left\{ \sum_{i \in I_n} \lambda_i x^i \mid \lambda \in S_n, \; x_i \in A (i \in I_n), n \in \mathbb{N} \right\}.$$

A set $U \subseteq \mathbb{R}^n$ is *located* if it is inhabited and if for all $x \in \mathbb{R}^n$ the distance

$$d(x, U) = \inf\{\|x - y\| \mid y \in U\}$$

exists, where throughout this paper $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$. U is said to be *totally bounded* if U is inhabited and if for every $\varepsilon > 0$ there exists a finite subset $F \subseteq U$ such that

$$\forall x \in U\ \exists y \in F\ \|x - y\| < \varepsilon.$$

Note that any totally bounded set is located [2, Proposition 2.2.9]. Let $x^1, \dots, x^n \in \mathbb{R}^m$ and recall that

i) $(x^i)_{i \in I_n}$ are *linearly independent* if

$$\forall\, \lambda \in \mathbb{R}^n\ (\|\lambda\| > 0\ \Rightarrow\ \|\sum_{i \in I_n} \lambda_i x^i\| > 0),$$

ii) $(x^i)_{i \in I_n}$ are *linearly dependent* if

$$\exists\, \lambda \in \mathbb{R}^n\ (\|\lambda\| > 0\ \wedge\ \sum_{i \in I_n} \lambda_i x^i = 0).$$

The following lemma seems to be folklore, but we could not find a proof in the constructive mathematics literature. As we will need it later on, we provide a proof for the sake of completeness.

Lemma 1. *Let $x^1, \dots, x^n \in \mathbb{R}^m$. If $n > m$, then $(x^i)_{i \in I_n}$ are not linearly independent.*

Proof. It suffices to prove the assertion for $n = m + 1$. Assume that $(x^i)_{i \in I_{m+1}}$ are linearly independent.
Case $m = 1$: By linear independence we have $|x^1| > 0$ and $|x^2| > 0$. Set $\lambda_1 := x^2$ and $\lambda_2 := -x^1$. Then $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2$ satisifes $\|\lambda\| > 0$ and

$$\lambda_1 x^1 + \lambda_2 x^2 = x^2 x^1 - x^1 x^2 = 0$$

which is a contradiction.
Case $m \geq 2$: As $\|x^{m+1}\| > 0$ we have that $|x^{m+1}_j| > 0$ for some $j \in I_m$. Without loss of generality we assume that $j = m$. Consider the vectors

$$v^i := x^{m+1}_m x^i - x^i_m x^{m+1}, \quad i \in I_m.$$

We have $v^i_m = 0$ for all $i \in I_m$, so we may identify the vectors v^i with elements of $\mathbb{R}^{m-1}$. Moreover, the $(v^i)_{i \in I_m}$ are linearly independent. Indeed, consider $\lambda \in \mathbb{R}^m$ with $\|\lambda\| > 0$, then

$$\sum_{i \in I_m} \lambda_i v^i = \sum_{i \in I_m} (\lambda_i x^{m+1}_m) x^i + (-\sum_{i \in I_m} \lambda_i x^i_m) x^{m+1}.$$

Since $|\lambda_k| > 0$ for some $k \in I_m$ and as $|x_m^{m+1}| > 0$ we have $\|\tilde{\lambda}\| > 0$ where $\tilde{\lambda} \in \mathbb{R}^{m+1}$ is given by $\tilde{\lambda}_i := \lambda_i x_m^{m+1}$, $i \in I_m$, and $\tilde{\lambda}_{m+1} = -\sum_{i \in I_m} \lambda_i x_m^i$. Linear independence of $(x^i)_{i \in I_{m+1}}$ now implies that

$$\| \sum_{i \in I_m} \lambda_i v^i \| = \| \sum_{i \in I_{m+1}} \tilde{\lambda}_i x^i \| > 0.$$

Thus, by erasing the last coordinate of the v^i, we have constructed m linear independent vectors in $\mathbb{R}^{m-1}$. Continuing this reduction procedure, if necessary, will eventually produce two linearly independent vectors in $\mathbb{R}$ which is a contradiction according to the case $m = 1$ above. □

Corollary 1. *Suppose that $x^1, \dots, x^n \in \mathbb{R}^m$ are linearly independent. Then*

(i) $n \leq m$;

(ii) if $n = m$, then $x^1, \dots, x^n$ is a basis of $\mathbb{R}^m$, that is

$$\mathbb{R}^m = \operatorname{span}(\{x^1, \dots, x^n\}).$$

Proof. (i) is obvious by Lemma 1. As for (ii), let $x \in \mathbb{R}^m$. Note that $V := \operatorname{span}(\{x^1, \dots, x^m\})$ is a closed located linear subspace of $\mathbb{R}^m$ ([2, Lemma 4.1.2 and Corollary 4.1.5]). We show that $\mathbb{R}^m \subseteq V$. To this end, let $x \in \mathbb{R}^m$. We have to show that $d(x, V) = 0$, that is $\neg d(x, V) > 0$. Assume $d(x, V) > 0$. Then $x^1, \dots, x^m, x$ are linearly independent, see [2, Lemma 4.1.10]. This is a contradiction to Lemma 1.

□

Lemma 2. *Let $x^1, \dots, x^n \in \mathbb{R}^m$. Then $\operatorname{co}(\{x^1, \dots, x^n\})$ is located. Moreover, if $n \geq 2$ and $x^1 - x^n, x^2 - x^n, \dots, x^{n-1} - x^n$ are linearly independent, then $\operatorname{co}(\{x^1, \dots, x^n\})$ is closed.*

Proof. Locatedness follows from [2, Propositions 2.2.6 and 2.2.9]. As for closedness, let $(y^k)_{k \in \mathbb{N}} \subseteq \operatorname{co}(\{x^1, \dots, x^n\})$ be a sequence converging to $y \in \mathbb{R}^m$. Further, let $\lambda^k \in S_n$ such that

$$y^k = \sum_{i=1}^{n} \lambda_i^k x^i = x^n + \sum_{i=1}^{n-1} \lambda_i^k (x^i - x^n).$$

Then

$$y^k - y^l = \sum_{i=1}^{n-1} (\lambda_i^k - \lambda_i^l)(x^i - x^n).$$

By linear independence of $x^1 - x^n, \dots, x^{n-1} - x^n$ the mapping

$$\mathbb{R}^{n-1} \ni \mu \mapsto \sum_{i=1}^{n-1} \mu_i (x^i - x^n)$$

and its inverse are bounded linear injections, see [2, Corollary 4.1.5]. Hence, the sequence $(\lambda_1^k, \dots, \lambda_{n-1}^k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^{n-1}$ is Cauchy and thus converges to $(\lambda_1, \dots, \lambda_{n-1}) \in \mathbb{R}^{n-1}$, and one verifies that

$$\lambda := (\lambda_1, \dots, \lambda_{n-1}, 1 - \sum_{i=1}^{n-1} \lambda_i) \in S_n$$

satisfies $y = \sum_{i=1}^n \lambda_i x^i \in \mathrm{co}(\{x^1, \dots, x^n\})$. □

Lemma 3. *For $n \geq 2$ fix $x^1, \dots, x^n \in \mathbb{R}^m$ such that $x^1 - x^n, \dots, x^{n-1} - x^n$ are linearly dependent. Moreover, let $x \in \mathrm{co}(\{x^1, \dots, x^n\})$. Then for each $\varepsilon > 0$ there exists $j \in I_n$ and $y \in \mathrm{co}(\{x^i \mid i \in I_n \setminus \{j\}\})$ such that $\|x - y\| < \varepsilon$.*

Proof. Let $\lambda \in S_n$ such that $x = \sum_{i\in I_n} \lambda_i x^i$, and let $M > 0$ such that $M > \|x^i\|$ for all $i \in I_n$. For all $i \in I_n$ either $\lambda_i > 0$ or $\lambda_i < \frac{\varepsilon}{2M}$. Suppose that there is $j \in I_n$ such that $\lambda_j < \frac{\varepsilon}{2M}$. Let $\mu_i := \lambda_i + \frac{\lambda_j}{n-1}$, $i \in I_n \setminus \{j\}$, and note that $\mu_i \geq 0$ for all $i \in I_n \setminus \{j\}$ and

$$\sum_{i\in I_n\setminus\{j\}} \mu_i = \sum_{i\in I_n} \lambda_i = 1.$$

Set

$$y := \sum_{I_n\setminus\{j\}} \mu_i x^i \in \mathrm{co}(\{x^i \mid i \in I_n \setminus \{j\}\}).$$

Then

$$\|x - y\| \leq \lambda_j \|x^j\| + \frac{\lambda_j}{n-1} \sum_{i\in I_n\setminus\{j\}} \|x^i\| \leq 2M\lambda_j < \varepsilon.$$

Hence, the assertion of the lemma is proved in this case. Thus we may from now on assume that $\lambda_i > 0$ for all $i \in I_n$. In that case, as $x^1 - x^n, \dots, x^{n-1} - x^n$ are linearly dependent, there is $\tilde{\nu} \in \mathbb{R}^{n-1}$ with $\|\tilde{\nu}\| > 0$ such that

$$\sum_{i\in I_{n-1}} \tilde{\nu}_i (x^i - x^n) = 0.$$

Let $\nu_i := \tilde{\nu}_i$ for $i \in I_{n-1}$ and $\nu_n := -\sum_{i\in I_{n-1}} \tilde{\nu}_i$ so that $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{R}^n$ satisfies

$$\|\nu\| > 0, \quad \sum_{i\in I_n} \nu_i = 0, \quad \text{and} \quad \sum_{i\in I_n} \nu_i x^i = 0.$$

In particular there exists $k \in I_n$ such that $\nu_k > 0$. Let

$$\beta := \max\left\{\frac{\nu_i}{\lambda_i} \mid i \in I_n\right\}.$$

Then $\beta > 0$ and for all $i \in I_n$ we have that $\tilde{\mu}_i := \lambda_i - \frac{1}{\beta}\nu_i \geq 0$ and

$$\sum_{i\in I_n} \tilde{\mu}_i = \sum_{i\in I_n} \lambda_i = 1 \quad \text{and} \quad x = \sum_{i\in I_n} \tilde{\mu}_i x^i.$$

Pick $j \in I_n$ such that $\nu_j > 0$ and

$$\beta - \frac{\nu_j}{\lambda_j} < \frac{\varepsilon\beta}{2M}.$$

Then $\tilde{\mu}_j < \frac{\varepsilon}{2M}$, so we are in the situation we covered in the first part of this proof and may thus construct $y \in \mathrm{co}(\{x^i \mid i \in I_n \setminus \{j\}\})$ such that $\|x - y\| < \varepsilon$. □

For the following lemma we recall that a subset M of a set N is said to be detachable from N if

$$\forall x \in N\ (x \in M \vee x \notin M).$$

Lemma 4. *Let $n \geq 2$ and $x^1, \ldots, x^n \in \mathbb{R}^m$. Suppose that the set*

$$\mathcal{L} := \{J \subseteq I_n \mid |J| \geq 2 \wedge \exists i \in J\ (x^j - x^i)_{j\in J\setminus\{i\}} \text{ are linearly independent}\}$$

is detachable from $\mathcal{P}(I_n)$. Then for all inhabited $J \subseteq I_n$ with $|J| \geq 2$ we have either

i) there exists $i \in J$ such that $(x^j - x^i)_{j\in J\setminus\{i\}}$ are linearly independent, or

ii) there exists $i \in J$ such that $(x^j - x^i)_{j\in J\setminus\{i\}}$ are linearly dependent.

Proof. Let $J \subseteq I_n$ be inhabited with $|J| \geq 2$. Note that

$$\{i, j\} \in \mathcal{L} \ \Leftrightarrow\ \|x^i - x^j\| > 0 \text{ and } \neg(\{i, j\} \in \mathcal{L}) \Leftrightarrow \|x^i - x^j\| = 0. \quad (1)$$

Hence, as $\mathcal{L}$ is detachable from $\mathcal{P}(I_n)$, for arbitrary $i, j \in J$ we have either $\|x^j - x^i\| > 0$ or $\|x^j - x^i\| = 0$, and thus we know whether there is $i, j \in J$ such that $\|x^j - x^i\| = 0$ or whether $\|x^j - x^i\| > 0$ for all $i, j \in J$. In the first case ii) holds. In the second, the set

$$\mathcal{L}(J) := \{J' \mid (J' \in \mathcal{L}) \wedge (J' \subseteq J)\},$$

which is detachable from $\mathcal{P}(I_n)$, is inhabited. Pick a set $\tilde{J} \in \mathcal{L}(J)$ of maximal cardinality. If $\tilde{J} = J$, then i) holds. If $\tilde{J} \subsetneq J$, let $i \in \tilde{J}$ such that $(x^j - x^i)_{j \in \tilde{J} \setminus \{i\}}$ are linearly independent. Note that $\operatorname{span}(\{x^j - x^i \mid j \in \tilde{J} \setminus \{i\}\})$ is located and closed ([2, Lemma 4.1.2 and Corollary 4.1.5]). For $k \in J \setminus \tilde{J}$ suppose that

$$d(x^k - x^i, \operatorname{span}(\{x^j - x^i \mid j \in \tilde{J} \setminus \{i\}\})) > 0.$$

Then $x^k - x^i, (x^j - x^i)_{j \in \tilde{J} \setminus \{i\}}$ are linearly independent ([2, Lemma 4.1.10]). Thus $\tilde{J} \cup \{k\} \in \mathcal{L}(J)$ which contradicts maximality of $\tilde{J}$. Hence,

$$d(x^k - x^i, \operatorname{span}(\{x^j - x^i \mid j \in \tilde{J} \setminus \{i\}\})) = 0,$$

that is ii) holds. □

Definition. *A formula φ is* conditionally constructive *if there exists a $k \in \mathbb{N}$ and a subset M of I_k such that the detachability of M from I_k implies φ.*

One verifies that conditionally constructive formulas are closed under conjunction and implication and may be used unconditionally in the proof of falsum:

Lemma 5. *Let the formulas φ and ψ be conditionally constructive. Then*

i) if $\varphi \Rightarrow \nu$, then ν is conditionally constructive,

ii) $\varphi \wedge \psi$ is conditionally constructive,

iii) $(\varphi \Rightarrow \neg\psi) \Rightarrow \neg\psi$.

Proof. See [1]. □

The following proposition shows that Carathéordory's Convexity Theorem is conditionally constructive.

Proposition 1. *Fix an inhabited set $A \subseteq \mathbb{R}^m$ and $x \in \operatorname{co}(A)$. Then the following statement is conditionally constructive:*

CCT(A) *There are vectors $z^1, \dots, z^k \in A$ with $k \leq m + 1$ such that*

$$x \in \operatorname{co}(\{z^1, \dots, z^k\}).$$

Proof. Let $x^1, \ldots, x^n \in A$ and $\lambda \in S_n$ such that $x = \sum_{i \in I_n} \lambda_i x^i$, and define $\mathcal{L}$ as in Lemma 4. Furthermore, define subsets $\Omega_i \subseteq \mathcal{P}(I_n) \times I_3$, $i \in I_3$, by

$$\begin{aligned}
(J,1) \in \Omega_1 &\Leftrightarrow\ J \in \mathcal{L}, \\
(J,2) \in \Omega_2 &\Leftrightarrow\ |J| \geq 1 \wedge d(x, \mathrm{co}(\{x^j \mid j \in J\})) = 0, \\
(J,3) \in \Omega_3 &\Leftrightarrow\ |J| \geq 1\ \wedge\ d(x, \mathrm{co}(\{x^j \mid j \in J\})) > 0.
\end{aligned}$$

Suppose that $\bigcup_{i \in I_3} \Omega_i$ is detachable from $\mathcal{P}(I_n) \times I_3$ which in particular implies that $\mathcal{L}$ is detachable from $\mathcal{P}(I_n)$. We prove, under this assumption, that there is $\tilde{J} \in \mathcal{P}(I_n)$ with $|\tilde{J}| \leq m+1$ such that

$$x \in \mathrm{co}(\{x^j \mid j \in \tilde{J}\}).$$

Suppose that $\Omega_3 = \emptyset$. Then, as $(\{j\}, 2) \in \Omega_2$ for arbitrary $j \in I_n$, we have in fact that $x = x^1 = \ldots = x^n$, and the assertion holds. Thus we may from now on assume that Ω_3 is inhabited. Let $\varepsilon > 0$ satisfy

$$\varepsilon < \min\{d(x, \mathrm{co}(\{x^j \mid j \in J\})) \mid (J,3) \in \Omega_3\}.$$

Note that Ω_2 is inhabited, because $(I_n, 2) \in \Omega_2$. Let $\tilde{J} \in \mathcal{P}(I_n)$ be of minimal cardinality amongst all $J \in \mathcal{P}(I_n)$ such that $(J,2) \in \Omega_2$. If $\tilde{J} = \{j\}$, then $x = x^j$, and the assertion is proved. Hence, we may assume that $|\tilde{J}| \geq 2$. By Lemma 4 either $\tilde{J} \in \mathcal{L}$ or there is $i \in \tilde{J}$ such that $(x^j - x^i)_{j \in \tilde{J} \setminus \{i\}}$ are linearly dependent. Suppose that latter, and let $y \in \mathrm{co}(\{x^j \mid j \in \tilde{J}\})$ such that $\|x - y\| < \varepsilon/2$. By Lemma 3 there is $k \in \tilde{J}$ such that

$$d(y, \mathrm{co}(\{x^j \mid j \in \tilde{J} \setminus \{k\}\})) < \varepsilon/2$$

which implies

$$\begin{aligned}
d(x, \mathrm{co}(\{x^j \mid j \in \tilde{J} \setminus \{k\}\})) &\leq \|x - y\| + d(y, \mathrm{co}(\{x^j \mid j \in \tilde{J} \setminus \{k\}\})) \\
&< \varepsilon.
\end{aligned}$$

Thus $\neg(\tilde{J} \setminus \{k\}, 3) \in \Omega_3$, that is $(\tilde{J} \setminus \{k\}, 2) \in \Omega_2$ which contradicts minimality of $\tilde{J}$. Hence, $\tilde{J} \in \mathcal{L}$. But then $\mathrm{co}(\{x^j \mid j \in \tilde{J}\})$ is closed by Lemma 2, so $x \in \mathrm{co}(\{x^j \mid j \in \tilde{J}\})$ follows, and also $|\tilde{J}| \leq m+1$ by Corollary 1.

□

As a consequence of Proposition 1 we obtain the already advertised approximate version of Carathéordory's Convexity Theorem for totally bounded sets, namely that the convex hull $\mathrm{co}(A)$ of a totally bounded set $A \subseteq \mathbb{R}^m$ is approximated up to arbitrary small error by the subset consisting of a all convex combinations of degree $m+1$:

$$\mathrm{co}^{m+1}(A) := \left\{ \sum_{i \in I_{m+1}} \lambda_i z^i \mid z^i \in A (i = 1, \ldots, m+1), \lambda \in S_{m+1} \right\}.$$

So if we could prove that $\mathrm{co}^{m+1}(A)$ and $\mathrm{co}(A)$ are closed, which we in general cannot, then $\mathrm{co}(A) = \mathrm{co}^{m+1}(A)$ as is classically always the case. Indeed classically $\mathrm{co}(A) = \mathrm{co}^{m+1}(A)$ is compact whenever A is compact.

Theorem 1. *Suppose that $A \subseteq \mathbb{R}^m$ is totally bounded. Then for every $x \in \mathrm{co}(A)$ and every $\varepsilon > 0$ there is $y \in \mathrm{co}^{m+1}(A) \subseteq \mathrm{co}(A)$ such that $\|x - y\| < \varepsilon$. In particular, $\overline{\mathrm{co}}(A) = \overline{\mathrm{co}^{m+1}}(A)$ where $\overline{\mathrm{co}}(A)$ denotes the closure of $\mathrm{co}(A)$ and $\overline{\mathrm{co}^{m+1}}(A)$ the closure of $\mathrm{co}^{m+1}(A)$, and $\overline{\mathrm{co}}(A)$ is compact.*

Proof. Let

$$\begin{array}{rccc} \kappa : & S_{m+1} \times A^{m+1} & \to & \mathbb{R}^m \\ & (\lambda_1, \ldots, \lambda_{m+1}, z^1, \ldots, z^{m+1}) & \mapsto & \sum_{i \in I_{m+1}} \lambda_i z^i. \end{array}$$

As κ is uniformly continuous and its domain is totally bounded, its range $\mathrm{co}^{m+1}(A)$ is totally bounded as well, see [2, Proposition 2.2.6], and hence $\overline{\mathrm{co}^{m+1}}(A)$ is compact. We show that $\mathrm{co}(A) \subseteq \overline{\mathrm{co}^{m+1}}(A)$. Fix $x \in \mathrm{co}(A)$. We have to show that

$$d(x, \mathrm{co}^{m+1}(A)) = 0,$$

that is

$$\neg(d(x, \mathrm{co}^{m+1}(A)) > 0).$$

According to Lemma 5 it suffices to prove this under the assumption that $\mathrm{CCT}(A)$ holds. But obviously

$$d(x, \mathrm{co}^{m+1}(A)) > 0$$

contradicts $\mathrm{CCT}(A)$. □

Note that the fact that the convex hull of a totally bounded set A is totally bounded, and thus its closure compact, is also easily directly verified. The important message of Theorem 1 is that the convex hull of A is best approximated by $\mathrm{co}^{m+1}(A)$. An inspection of the proof shows that we could replace the requirement of A being totally bounded in Theorem 1 by $\mathrm{co}^{m+1}(A)$ being located which, however, does not seem a very useful generalisation.

References

[1] Josef Berger and Gregor Svindland. On Farkas' lemma and related propositions in BISH. *preprint*, 2021. https://arxiv.org/abs/2101.03424.

[2] Douglas Bridges and Luminita Vîţă. *Techniques of Constructive Analysis*. Springer, 2006.