Review of "Minkowski-Weyl Theorem", a bachelor thesis of Nathan Chappell

Summary of Results

The first part of the thesis is a proof of the Minkowski-Weyl Theorem. The proof is first done in the special case of cones, and then a reduction from the case of general polyhedra to the case of cones is described. The main tool here is Fourier-Motzkin elimination, hence the proof is constructive, and the author has implemented it in C++.

The second part of the thesis discusses the question of how one might test the validity of the output of the implemented program, and touches on the notions of full-dimensionality, pointedness, and the Farkas lemma. The developed equivalence criteria are also subsequently implemented.

Evaluation

Strengths

Non-trivial theory: the thesis discusses non-trivial theoretical results. It is difficult for me to say how much of the second part of the thesis is original, but regardless the thesis testifies to the fact that the author has grappled well with fundamental notions of polyhedral theory.

C++ implementation: it is notoriously hard to implement some "simple" proofs, so I applaud the authors for taking up this endeavour. (Diclaimer: I don't code in C++, so I can't judge the quality of the code, but it looked good to me.)

Hepful figures: The thesis contains several helpful figures and diagrams.

Weaknesses

The main weakness of the thesis is the quality of write-up, which suffers from the following:

- **improper or missing quantification:** a majority of statements is quantified implicitly from the context, to the point that the reader has to figure out dimensions of vectors and matrices, or even whether a statement is quantified existentially or universally. This has significant impact on the readability: in the second part (Section 4) it was often not (immediatelly) clear to me what is assumed and what is tested.
- **backward definitions:** often a notion or notation is first used and only afterwards defined.
- **matrices or sets of vectors:** the author treats matrices as sets of vectors, often defining a matrix as a set, performing set difference on it, etc., but also performing

matrix multiplication with it. This leads to strangeness such as $\exists \mathbf{v} \in V : \mathbf{v} = V\mathbf{e}_i$ which is just a complicated way of saying that $\mathbf{v}$ is the i-th column of V...

- **boxes and frames:** the author uses two different shades to distinguish grey boxes for definitions and propositions, and uses frames for theorems and remarks. The frames are ugly, but the boxes are actually fairly nice, so this is a neutral point.

Detailed Comments

In order to give detailed comments, I'm attaching a version of the thesis with added line numbers. These numbers are not quite accurate, but are sufficient for this purpose. The coordinates below are in the (page,line) format, where the page numbers are those in the attached document (which are shifted with respect to the original thesis)

- (3,6) don't use numerical references in an abstract, rather give the full name – here, it would be something like "Ziegler [Lectures on Polytopes, Springer]"
- (5,6) here and also in that section: "Characterstic" $\Rightarrow$ "Characteristic"
- (7,5) you should first say that U^j is a column. But this is a general comment: you often use notation that you only define later or even assume to be standard. I suggest, especially in the context of a bachelor thesis, to err on the side of defining too much.
- (7,13-15) why is this note not a definition of the Minkowski sum?
- (9,0) rephrase the quantification as "and let $\{\mathbf{x}^i \mid i \in [n]\} \subseteq C$, and $c_i \geq 0$ for each $i \in [n]$. Then: ..."
- (9,4 and 7): the "So, ..." sentences are redundant.
- (9,4) if you really wanted to include the "So, ..." sentence, then "Next,..." should be a new paragraph.
- (9,5) why do we have $\mathbf{x}^i$ but $\mathbf{t}_i$ and c_i – a mismatch in sub- and superscripts?
- (9,9) the word "that" would be better replaced with "this fact"
- (9,20) "objects" $\Rightarrow$ "object"
- (9, Notational Abuses) Why do you use $\langle \mathbf{x}, \mathbf{y} \rangle$ for dot product when dot product is really a special form of matrix product and it would be less notation to just write $\mathbf{x}^\intercal\mathbf{y}$ or even simpler $\mathbf{xy}$? (This is frowned upon by some, but also used quite widely.) Either way, dealing with $\langle\rangle$ seems unnecessary.
- (10,6) why is the theorem in a frame identical to the one for remarks? Also, the frames are not very nice with the unevenly thick vertical and horizontal lines, and generally considered to be a typographically suboptimal means of highlighting. My suggestion is to use shaded backgrounds for theorems and definitions, and no background for remarks and theorems.
- (10,13) a backward definition: you first mention the term "coordinate projection" and only later define it. This is not good exposition, and Def 2.1.1 should be shifted before the narrative.
- (10,18-21) There's lots of redundance. If the proof is simple, demonstrate it by the proof being short – there's no need to tell me that you will use Lemmas 2.1.2 and 2.1.3 – just use them!

- (10,27) the phrase “takes one step” (and later “take the second step”) is unclear and confusing, please rephrase.
- (10,29) “equality” should be “equivalence” (these two are not the same)
- (10,32) the word “comparing” seems off, perhaps “combining” or something else?
- (11,3) the description of Π is correct but somewhat clumsy, I suggest writing $(I\ \mathbf{0}) \in \mathbb{R}^{d\times(d+p)}$, here and later (the phrasing repeats).
- (11,15) replace by “Since this holds for every row B_i”
- (11,19) you don’t quantify k and you should
- (11,19) you have not defined the unit vectors I believe
- (11,19) what is the role of saying $d+p$? For the purposes of this Proposition, all that matters is that B and B' have the same number of columns which is typically called n. d and p are parameters from the ongoing construction, but it is good writing to state things generally.
- (12,1-3) my guess is that your TeX code here used $\{\}$ instead of \{\}. The result is quite confusing, because it seems like you’re defining three numbers P, N and Z several times. (This mistake repeats.)
- (12,7) while the use of italics makes it fairly clear that you are refering to the points of the statement, it would be better to say “Points 1 and 2 are clear.”
- (13,25-29) consider using the algorithm2e environment.
- (13,27) capital P in “proposition” (here and generally)
- (14,0-13) while this seems all correct, I think it might be clearer to phrase this as a construction of a series of cones of decreasing dimensions whose restriction to the original coordinates stays the same. (The notion at play is that of an “extended formulation”, which might also be helpful to define.) The current exposition seems to somewhat obscure what is going on geometrically, but this is a matter of taste.
- (14,28-29) again a backward definition of U'
- (15,5-6) redundant sentence (if it’s simple, just state it)
- (15,11) A small comment on the use of $\leq$ in the definition of a cone: while this is correct and what Ziegler uses, it would seem more intuitive to me to use $\geq$, since $I\mathbf{x} \geq \mathbf{0}$ defines the non-negative orthant which feels more fundamental, while $I\mathbf{x} \leq \mathbf{0}$ defines the non-positive orthant. (The definition with $\geq$ is used on wikipedia, for example.)
- (15,16) is $T_H(A)$ a collection of vectors or a matrix? The distinction is that a matrix has an order, and this order matters (e.g., which column/row you eliminate with F-M). Here and elsewhere you often treat matrices as sets of vectors, and sometimes they are column vectors when other times they are row vectors. I strongly suggest being less sloppy here even at the cost of more verbosity.
- (16,0-9) this whole description is a very literal (non-intuitive) way of saying “split variables into positive and negative and add slack variables”. These are fairly well-known notions and it would be helpful to state them explicitly before giving the details.
- (16,20) here you quantify k – this should be replicated in the previous version.
- (17,1-3) same as (13,1-3)
- (17,5) same as (15,16)

- (17,18) “by THE Closure Property of Cones” (missing article), plus add a reference to the Proposition
- (17,21) here and elsewhere: when a displayed equation is a continuation of a sentence and the sentence ends here, the equation should contain a period separated by an enspace, so your TeX code (of the equation) should contain “\enspace .”
- (17,27) replace “in this equality” by “value of the last term”
- (18,21) define Π first
- (18,21) you should write $\{(\mathbf{x}, x_0) \in \mathbb{R}^{\cdots} \mid [-\mathbf{b}|A](x_0, \mathbf{x}))\}$ – the present “definition” does not make any sense.
- (19,0-28) it is confusing that the columns of U' are denoted U^i, since one would expect these to be columns of U.
- (19,11) I can’t parse this: how could the first cone be contained in the second cone when the first cone has dimension one more than the second cone?
- (19,28) don’t use the phrase “fairly clear” in a proof. If it’s clear, your job is done. If not, you still have work to do.
- (20,9-15) good job! I like this diagram. My one comment is that above the bidirectional arrows, it is not clear which label belongs to which direction. I suggest using bidirectional harpoon style arrows as discussed here [1].
- (21,9) “transform IS sent to stdout”
- (21,32) “points” $\Rightarrow$ “vertices” (here and elsewhere)
- (30,2) use the TeX command “\\” to force a line break before “generalized_lift” to avoid line overfull
- (32,9) here you use “:” in set notation, but previously you have used “|”. It is best to use the macro “\mid” and let the style file define this either way. It doesn’t matter too much which you use, but it is necessary to be consistent.
- (35,13-16) here and elsewhere you provide such lists of implications. In particular I don’t understand the meaning of the second line. In general I recommend being much more formal: saying what is given, what is computed by the program, what is tested, what follows from what.
- (35,24) here and elsewhere when you mean strict containment I suggest using $\subsetneq$ which highlights the fact that you want to rule out equality. (Some authors use $\subset$ as a synonym for $\subseteq$.)
- (35,25) here and elsewhere you use $\mathbf{v} = V\mathbf{e}_i$ to say that v is the i-th element of V. This would not be necessary if you strictly treated V as a matrix (just like you do when you write $\mathbf{v} = V\mathbf{t}$ on the same line) and you could refer to its rows and columns.
- (36,13) again, “THE Closure Property ...”, and give a reference to the Proposition (by its number)
- (36,20) typo $\mathbf{t}_{1,2}$ should be $\mathbf{t}_1, \mathbf{t}_2$, right?
- (37,1) this is particularly bad example of not quantifying your statement: is this $\exists\mathbf{t}$ or $\forall\mathbf{t}$? (I know which it is, but you should not make the reader guess.)
- (37,9) wont?
- (37,19) “there exist l, t” (plural)
- (38,8) “we run the program” - on what?

- (38,18) here and elsewhere: “criteria” is the plural of “criterion”. You could mean that one has to check two things to verify the equivalence, so “Criteria 1” stand for “the first set of criteria”, but it seems to me to make more sense to just say “Equivalence Criterion 1” (where the criterion is a conjunction of things that have to hold)
- (38,21) here and elsehwhere: what does it mean to “hand-craft” something? Does it mean that you construct a minimal V and also construct a matrix A such that they correspond, and then you verify A using the knowledge of V? As I said above, it wasn't very clear what are the assumptions here.
- (38,33) Fourier-Motzkin (capital letters – these are names)
- (39,17) again, it would be very helpful to quantify precisely, and to write the dimensions of the discussed vectors.
- page 40: having **x** have dimension $d + n$ for a part of the proof and then suddenly dimension d is very bad.
- (40,17) backward definition of Y'
- (40,18) why the parentheses around (U)?
- (41,16) ambient space has not been defined
- (41,22) “for every **x** IN the cone”
- (42,26) refer to the Proposition by its number
- (42,31) as far as I can see V' is undefined.
- (43,0) is H supposed to be A? How is A' quantified? Plus, I think you want to drop “generating” (hyperplanes are not generating, or you haven't said so – you have defined a minimal set of hyperplanes). Same for (43,16)
- (45,15-16) “Characterstic” $\Rightarrow$ “Characteristic”
- (45,17) lower case “p” in “polyhedron”
- (45,18) quantify r in the assumptions of the Proposition
- (46,7) probably again you need to use “\{\}” to get braces
- (49, Thm 5) quantify your assumptions
- (49,14) you should almost never use “you” in formal writing. Here it would be “we get the proposition.”
- (49,24) this definition seems strange and non-intuitive. It would be better to say that a set of vectors V is full-dimensional if conv(V) contains a full-dimensional ball of some small radius, and then derive your desired property from this (more intuitive) definition, OR explain your definition intuitively.
- (50,14) delete “Then”
- (50,22) quantify A, A', **b**, **b**$'$
- (52,11) something weird is going on typographically: P' is closer to the arrow than the letter above and below it
- (53,19) lowercase “c” in “Cone”
- (53,31-37) change the tone to avoid using “you”
- (61,31-36) you say that the double description method is more efficient. While this may be true, some blow-ups are unavoidable, as the hypercube in n dimensions has 2^n vertices but $2n$ facets, and the cross-polytope (the dual of the hypercube) has n vertices but 2^n facets. So it would be helpful to clarify in what way is the double

description method more efficient.

[1] https://tex.stackexchange.com/questions/387921/drawing-bidirectional-arrows-in-a-latex-graph

Conclusion

The author has more than sufficiently demonstrated his ability to deal with non-trivial notions in polyhedral geometry and even pursue new questions. The C++ implementation constitutes another piece of evidence that his understanding is solid. However, because it is equally important to be able to communicate results well and clearly, and the presented work contains a decent amount of minor and not so minor write-up issues, I recommend the thesis be accepted as a bachelor thesis with the grade "2".

In Prague, August 23th, 2019