



HAL
open science

What are the most used guitar positions?

Jules Cournut, Louis Bigo, Mathieu Giraud, Nicolas Martin, David Régnier

► **To cite this version:**

Jules Cournut, Louis Bigo, Mathieu Giraud, Nicolas Martin, David Régnier. What are the most used guitar positions?. International Conference on Digital Libraries for Musicology (DLfM 2021), 2021, Online, United Kingdom. pp.84-92, 10.1145/3469013.3469024 . hal-03279863

HAL Id: hal-03279863

<https://hal.science/hal-03279863>

Submitted on 29 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

What are the most used guitar positions?

Jules Cournut

Louis Bigo

Mathieu Giraud

Univ. Lille, CNRS, Centrale Lille,

UMR 9189 CRISTAL

Lille, France

Nicolas Martin

Arobas Music

Lille, France

David Régnier

Univ. Lille, CNRS, Centrale Lille,

UMR 9189 CRISTAL

Lille, France

ABSTRACT

The playing of stringed instruments can be notated with symbolic *tablatures*, indicating for each note, the string, and the position on the fretboard at which it has to be played. We propose an encoding for tablatures in which the *hand position on the fretboard* is encoded to ease computational analysis and generation of guitar music. We analyze a corpus of 1022 guitar tablatures in standard tuning that were accurately transcribed in the corpus *MySongBook*, spanning metal, rock, pop, jazz/blues, and baroque/classical/romantic repertoires. We show how some of the proposed encodings reveal relevant statistics and some invariance properties on such guitar music. We show study unusual guitarist gestures around the G/B string pair due to its interval of major third, compared to the perfect fourth between other adjacent pairs of strings. We study the corpus across genres and textures (rhythm and lead parts) and how same relative chord positions are used along the fretboard. We distribute sets of 1000 most frequent vectors in different encodings for these five music genres.

CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; • **Information systems** → **Music retrieval**; *Data encoding and canonicalization*.

KEYWORDS

corpus, guitar tablature, quasi-isomorphism, hand position, chords

ACM Reference Format:

Jules Cournut, Louis Bigo, Mathieu Giraud, Nicolas Martin, and David Régnier. 2021. What are the most used guitar positions?. In *8th International Conference on Digital Libraries for Musicology (DLfM2021)*, July 28–30, 2021, Virtual Conference, GA, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3469013.3469024>

1 INTRODUCTION

Manually plucked stringed instruments were already present about 5,000 years ago in Egypt and then more widely around the Mediterranean. In Europe and the Middle East, there were several types

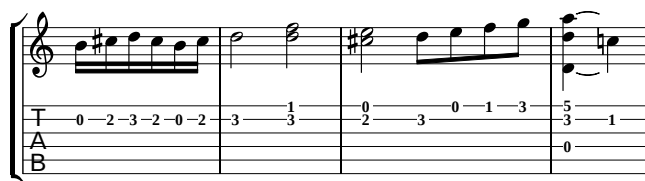
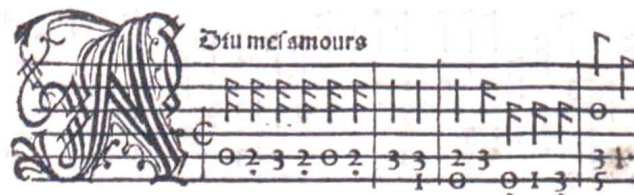
Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

DLfM2021, July 28–30, 2021, VirtualConference, GA, USA

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8429-2/21/07...\$15.00

<https://doi.org/10.1145/3469013.3469024>



```
E -----1-----0-----0--1--3--5-----  
B --0-2-3-2-0-2-3--3--2--3-----3--1---  
G -----  
D -----0-----  
A -----  
E -----
```

Figure 1: Francesco Spinacino lute intabulation of Josquin Desprez polyphonic chanson *Adu mes amours*, printed in 1507 (top, RISM 993103811, Biblioteka Jagiellońska, Kraków, P 680), with a possible transcription (middle) and an ASCII tab (bottom) on a modern guitar tablature assuming a standard tuning.

of *lutes*¹ in the Middle Ages, and the *vihuela* in Spain, in 15th and 16th centuries. The spanish word *guiterna* was used as early as in the 14th century, but the ancestors of guitars as we know them appeared in the 16th century. The current form of the “classical guitar” comes from the Spanish guitar maker Antonio de Torres in the 19th century. The first electric guitars date from the beginning of the 20th century. The sound is still being produced by acoustic vibrations, but they are now amplified by a pickup.

Classical and then electric guitars were popularized through their use in jazz, rock and other contemporary music repertoires. Today, the guitar is one of the most played instruments, especially among

¹not to be confused with the *lute family*, family 321 of the Hornbostel-Sachs classification [24], which gathers all instruments where one or more strings are fixed in a plane parallel to a sounding board, and therefore including string instruments that are bowed, stricken or plucked.

the amateur musicians, with 2 million guitars sold annually in the United States [22] and probably more than 200 million guitarists worldwide.

All of these instruments include several strings, each tuned to a specific pitch. Today most guitars have six strings tuned to pitches E^3 , A^3 , D^4 , G^4 , B^4 , and E^5 . As with any instrument, the very construction of the guitar influences the music that is played. This article focuses on *tablatures* that describe such playing.

1.1 Tablatures

A number of musical instruments offers to the performer different options for playing one same note. With keyboard instruments, one has to choose which finger to put on the key. On (multi-)stringed instruments, the musician can play most of the notes on different strings and in different positions: They therefore have to choose the string and the finger.

We today call *tablatures* forms of music notation that lay out some of these performance details, thus freeing the musician from some decisions. The first tablatures, designed to indicate fingers on notes played by keyboard instruments (organ, then harpsichord), probably appeared in the Middle Ages. Tablature notations were then mainly developed for the instruments with plucked strings – lute and viola da mano, and later guitars – indicating the string and the position on the neck rather than the pitch of the note.

With the emergence of musical printing, tablatures became quite widespread. The very first instrumental music that Ottaviano Petrucci printed in Venice in 1507 included a collection of lute tablatures – the *Intabolutura de lauto*, arrangements by Francesco Spinacino of vocal pieces [6]. Lute tablatures were then written from top to bottom, from the lowest-sounding to the highest-sounding, in the form of horizontal lines (Figure 1, top). The numbers written on these lines indicate at which *fret* the string has to be plucked to produce the desired note. Today, most guitar tablatures represent the strings from bottom to top (Figure 1, middle and bottom, and Figure 2, left column).

Representing music in tablature highlights the *quasi-isomorphism* of the guitar. The same pattern, melodic or harmonic, can be played:

- at different positions on the fretboard (when it does not include any open strings);
- or even on different strings. Indeed, with the standard tuning ($E^3 A^3 D^4 G^4 B^4 E^5$), there is a perfect fourth between two successive strings (except between the G^4 and B^4 strings, where there is a major third).

A tablature can also bear information about the musical style as some guitarists tend to favor selected positions over others [14]. Tablatures may also contain additional musical content in addition to pitch and frets, such as the prescriptive symbols for *bends*, *slides*, *hammer-on*, and *pull-off*, *palm mute*, and *let ring*. Such playing techniques largely contribute to musical style.

1.2 Research in Computer Music, Guitar Music, and Tablatures

Guitar music gives rise to various research questions in computer music and more specifically in MIR (Music Information Retrieval). A number of works deal with audio data, including effect detection [19], playing technique analysis [12, 20], solo detection [7, 16], and automatic audio transcription to tablature [11, 25, 27].

Research topics in tablature symbolic representations include automatic *fingering*, understood as *score to tablature*, that is determining the string/frets combinations, but not the actual selection of fingers. They use genetic algorithms [23] or machine learning techniques and representations such as Markov models [10] and *n*-grams [15]. Other works focus on tablature generation by assembling licks with integer programming [5] or generating patterns using Hidden Markov Models (HMM) [14]. Note that music generation systems not taking into account the construction of an instrument might create music that is not playable.

Guitar tablature data are also used in other MIR tasks like chord analysis [17] or guitar identification [3]. The singular pitch topology of the fretboard has inspired some more theoretical researches trying either to generalize this concept through a transformational approach [4] or to develop a dedicated notational system with Klumpenhouwer networks [21].

The guitar plays an essential role in pop/rock music, and a number of researches have therefore focused on this repertoire. Related to the present work, the principle of *affordance* and its impact on guitar performance and composition has been a particular subject of attention [18, 28].

1.3 Overview

How to efficiently represent and encode guitar tablatures for computer music analysis and generation? We study several encodings of guitar tablatures (Section 2) elaborated for machine learning manipulation, in particular by taking into account the position of the hand on the fretboard that we call *RootFret*. We provide a parser to load Guitar Pro GPF files as python music21 objects. By analyzing a corpus with more than 1000 pieces and 5 millions of notes (Section 3), we show that these encodings allow us to retrieve certain musical properties on notes and chords across genres: Encoding relatively tablatures highlights the quasi-isomorphisms on the guitar (Section 4). We therefore believe that such encoding methods using the hand position could be useful for music information retrieval tasks in analysis and generation on guitar music (Section 5).

2 TABLATURES AS BINARY VECTORS

The usual *piano-roll* encodes music into two dimensions, the first one representing the pitch of the notes, the second one their onset and durations. The state of a piano-roll at a given moment (for example at each sixteenth note) can be represented by a binary vector. Each component of the vector encodes a note that can be either played (1) or not (0) at this very moment. Such a vector can be called *many-hot*, where zero, one, or several components can be equal to 1, as opposed to *one-hot* binary vectors, where one and

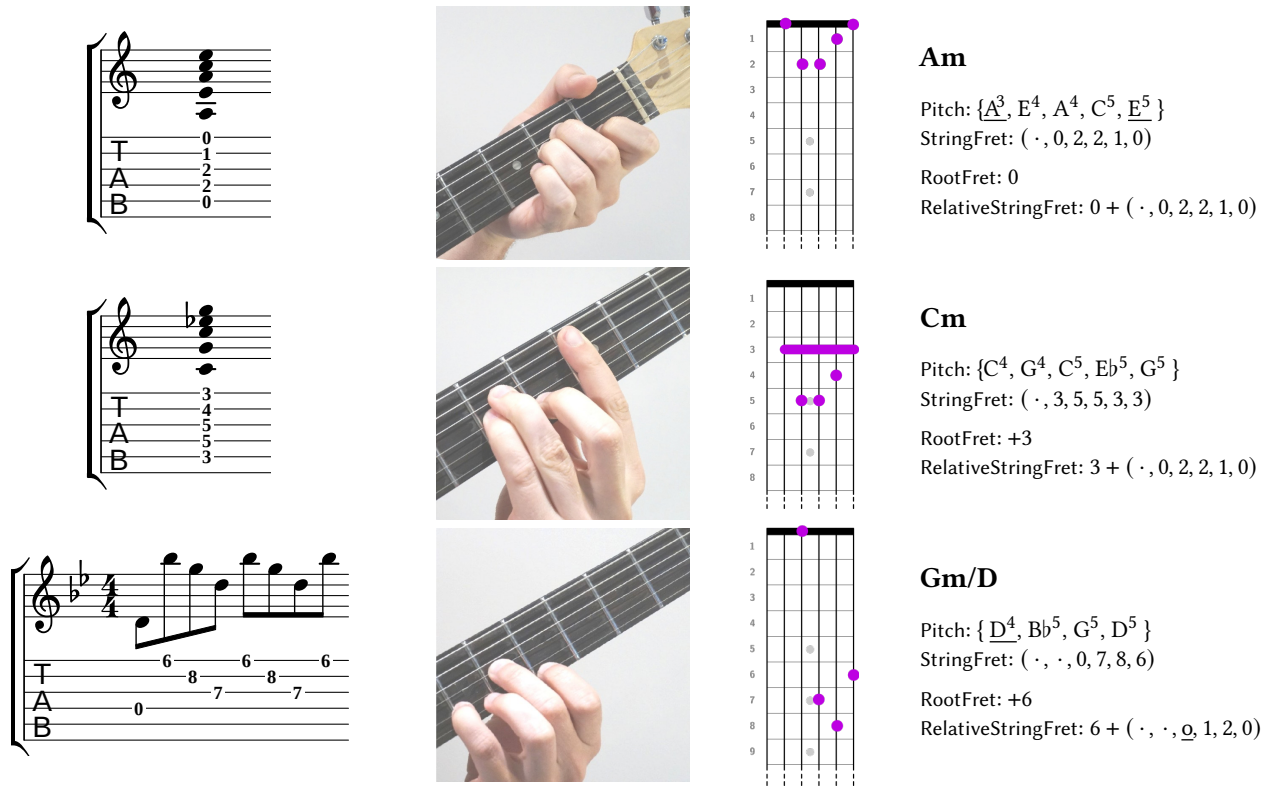


Figure 2: Tablatures, pictures, chord diagrams, and encoding of chords on a 6-string guitar with the standard tuning. (Top) The A minor chord uses two open strings (underlined, A^3 and E^5) and three fretted strings. The lowest string (\cdot) is not played. (Middle) The C minor chord, with the index finger used to bar the strings on the third fret, reproduces the A minor chord a minor third above, again without playing the lowest string. The `RelativeStringFret` encoding shows thus that both chords have the same relative position $(\cdot, 0, 2, 2, 1, 0)$. (Bottom) The broken chord of G minor with a D bass, in Scorpions’ *Still Loving You* (music: Rudolf Schenker), uses fretted strings starting from the sixth position of the fretboard, together with the open string D^4 . The `RelativeStringFret` encoding uses this high position (6) but still allows the open D string (\circ). The relative position is here the same as the one of the standard D minor chord $1 + (\cdot, \cdot, 0, 1, 2, 0)$.

only one component is equal to 1. Machine learning models are generally more efficient, and require less input data, when such vectors have a limited size and when the encoding of the input data is adapted to the problem.

Our motivation here is to propose guitar-specific encodings of tablature information. Ideally, an appropriate encoding should reflect positions that can be physically played and facilitate MIR tasks relating either to analysis or generation. In the following, we call `RootFret` an integer denoting the hand position on the fretboard – that is most of the time the position of the index finger, even if that finger is not playing at the very moment. A `RootFret` of +3 means thus playing with fingers starting at the third fret – what is usually called “third position” – possibly with a bar such as with the barred C minor chord (Figure 2, middle). As an exception, the “first position”, playing at the top of the fretboard, is denoted with a `RootFret` of 0 (and not +1), because in that first position open strings are only 1 semitone close to fretted strings (Figure 2, top).

For the sake of simplicity, this study concerns tablatures for guitars having $s = 6$ strings with standard tuning ($E^3, A^3, D^4, G^4,$

$B^4,$ and E^5) and that do not require the use of a capo. As the studied corpus includes a majority of tablatures in pop/rock style for electric guitar, we systematically consider the guitar neck to have $f = 24$ frets, which is not constraining for tablatures written for guitars with less frets.

Beyond the piano roll (Pitch), we study two additional encodings of guitar tablatures, denoting strings and frets, `StringFret` and `RelativeStringFret` (Figure 2) [1]:

- **Pitch.** The usual piano-roll (here with 49 notes, four-octave range of the guitar with the standard tuning, $E^3 - E^7$). This encoding does not indicate on which string a note is played, and is therefore not able to distinguish the playing of the same note on two different strings.
- **StringFret.** Six one-hot vectors of size $25 = f + 1$, representing, for each string, the 24 frets plus the *open string*.
- **RelativeStringFret.** One one-hot vector with the hand position `RootFret` (size $25 = f + 1$), and six one-hot vectors of size $h + 1$ encoding the position relatively to the `RootFret`. The

	n	div.	C10	C100	unique
Pitch	49	8 485	22.8%	68.3%	583
StringFret	150	11 429	14.5%	57.5%	686
RelativeStringFret	61	12 616	10.7%	43.2%	826
RelativeStringFret*	36	4 581	35%	81%	268

Table 1: Statistics on the 2.4 million non-empty vectors (without sustain information) from the MySongBook corpus. Div: Number of distinct vectors among the corpus. C10/C100: Cumulative sum of the number of occurrences of the 10/100 most frequent vectors. Unique: Vectors appearing only once in the corpus. RelativeStringFret* shows statistics on the vectors while ignoring the components encoding the hand position.

value h is the maximum “hand span”, and the +1 component encodes (not relatively) the open string, that can be used even in higher positions on the fretboard, as with the Gm/D chord on the bottom of Figure 2.

The latter encodings are more compact and more adapted to the instrument: compared to the 2^{49} theoretically possible vectors of the Pitch encodings, the spaces of StringFret (25^6 vectors) and RelativeStringFret ($25 \times (h+1)^6$ vectors) are much smaller. Note that only a small subset of the 25^6 combinations actually appears in guitar tablatures (see Section 4.1) as a majority of these positions implies to simultaneously press the fretboard at positions that are too far apart for a hand with an average size. With small values of h however, a significant proportion of the $(h+1)^6$ positions can actually be played.

The RelativeStringFret encoding accounts for the quasi-isomorphism between several positions along the fretboard (Figure 2). Any chord or pattern (such as a scale) without open string can thus be transposed by changing only the RootFret component.

Where is the hand on the neck? The guitar player generally minimizes the hand moves, having more dexterity with their fingers than with their wrist or their arm [8]. The RootFret value r is estimated here by a conservative heuristic: It has the value 0 at the beginning of the piece, and, at each note, minimally changes only when its current value r would not allow to play the next note or chord on the $[r, r+h]$ frets, h reflecting the size of the hand.

Sustained notes. Unlike encodings describing notes one by one, each one with its own duration [9], some piano-roll and derived encodings such as Pitch do not differentiate sustained notes from re-plucked notes. For Pitch, 49 additional components are needed, indicating whether a sounding note is sustained from the past or presently attacked. For StringFret and RelativeStringFret, only 6 additional components are needed because attack and sustain are valid information for an entire string.

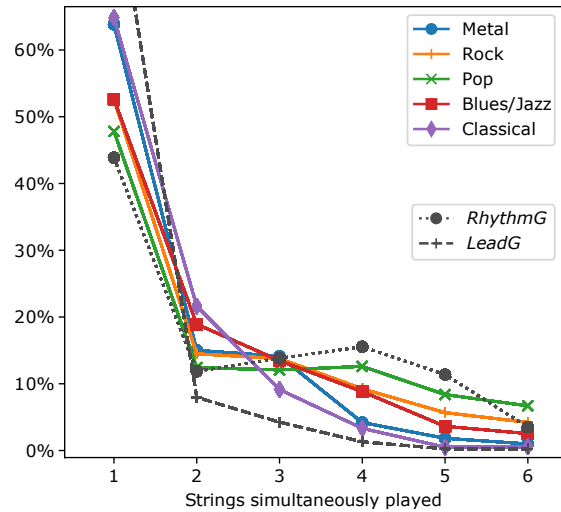


Figure 3: Number of simultaneously played strings according to the genre and the role of the guitar. More than 80% of vectors tagged with LeadG encode a single note.

3 CORPUS

3.1 Formats for Tablatures

Many guitar players share tablatures. Tablature formats are usually dedicated to sight reading or to music software usage. The popular ASCII tabs represent tablatures using single characters such as dashes or numbers (Figure 1, bottom) [26]. More structured formats are the .gp format, produced by the company Arobas Music for the proprietary Guitar Pro software, and the .tg format for the open-source software TuxGuitar. These formats allow to represent not only the notes and frets, but also a wide range of playing techniques as those mentioned in page 85.

The .gp format is not closed-source, but is currently not documented. However, it can be imported by most tablature edition software, including Ultimate Guitar, Songsterr, Soundslice, MuseScore, TuxGuitar, and Notion. A .gp file is an archive mainly containing a text file score.gpif, representing in XML pitch and fret information, but also playing techniques, for one or several instruments, with also the volume and associated audio effects, as well as an annotation of the overall structure of the part.

We developed a parser for such .gp files in the Python music21 library [2], available from www.algomus.fr/code. String and fret data are added to music21 streams by the attributes articulations.StringIndication and FretIndication.

3.2 The Corpus

The corpus MySongBook contains more than 2000 tablatures of different genres accurately transcribed by professional musicians² hired by Arobas Music. Each transcriber is specialized in some genres. The pieces are labelled with genre tags including rock, pop, blues/jazz, metal and classical. This classification is fairly rough, as styles and playing techniques may be different between sub-styles.

²<https://www.mysongbook.com/transcriber-team>

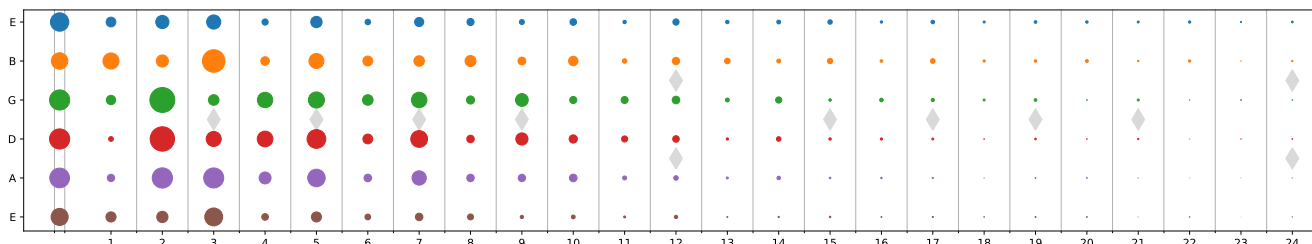


Figure 4: Strings and frets used on the 2.4 million vectors on the *MySongBook* corpus. The left column corresponds to open strings. Each dot has a size proportional to the number of vectors encoding that pitch.

This is the case for the blues/jazz class, or for the classical class that includes baroque, classical, and romantic tablatures.

The pieces have been selected based on their popularity within the guitarist community: The corpus is reasonably representative of what guitarists are generally playing, Arobas Music trying to satisfy its customers. Most pieces are original works for guitar, but some are arrangements. The transcriptions are intended to be highly accurate. For example, both in latin/jazz and pop styles, solos are transcribed note-by-note with extended information on the playing techniques. Transcribers largely refer to live video to favor original fingerings even if some other positions, possibly easier, might be found.

We focused on guitar tracks (including transcriptions of vocal parts for guitarists) with the standard tuning $E^3 A^3 D^4 G^4 B^4 E^5$, that is 2248 tracks in 1022 songs or pieces. Indeed, many pieces have several guitar tracks, with different tracks for the lead guitar(s) and the rhythm guitar(s). All of them are considered here. There are about 40% of rests in these tracks. The final corpus contains five million notes, sampled with the sixteenth note into 2,420,586 non-empty vectors. Other subdivisions could also be studied.

Role labels. Guitarists play different roles within a song. The *rhythm guitar* generally plays the harmony, sometimes with broken chords, or to accompany some lead instrumental or vocal part, whereas the *lead guitar* generally plays *riffs* and *solos*. We annotated measures of 169 tracks, corresponding to 50 songs and 12,891 non-empty vectors, as either *RhythmG* or *LeadG*, the latter also including some hybrid parts such as non-rhythm accompaniment parts. The *RhythmG* sections are prominent in the corpus, with 78% of the annotated measures.

3.3 Data Distribution

The raw corpus is not available due to legal constraints. However, we distribute in an open licence at www.algomus.fr/data the 1000 most frequent vectors in this corpus in all encodings and according to genre tags.

4 RESULTS AND DISCUSSION

This section shows how the proposed encodings and the corpus illustrate some properties of guitar music across different genres, including its quasi-isomorphism on the neck position and on the strings. After some generic statistics (Section 4.1), we focus on

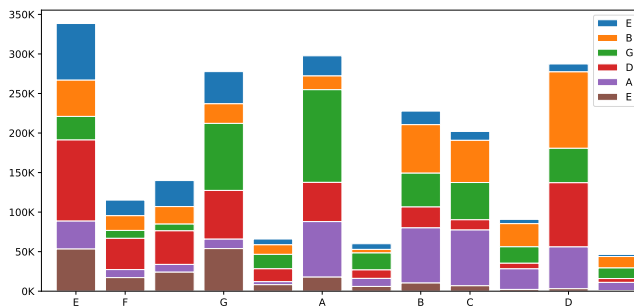


Figure 5: Pitches played, regardless to the octave, among the different strings on the 2.4 million vectors in the corpus.

successions of single notes (Section 4.2) and on absolute and relative position of chords (Section 4.3).

4.1 Vectors, Strings, and Frets

Vector Diversity. Table 1 shows how the proposed encodings behave over the 2.4 million vectors of the corpus. The *RelativeStringFret* encoding gathers more different vectors (12,616) than *StringFret* (11,429) because the same (absolute) position can sometimes be encoded with several different *RootFret* as its estimation depends on the past notes. However, when the hand position along the neck is ignored (*RelativeStringFret**), the number of unique vectors falls to 4 581, showing the ability of relative encodings to compress tablature information.

Sustained Notes, Number of Strings. There are 34.8% of the positions which contain vectors encoding sustained notes, indicating strings which were previously sounding. Only in less than 2% of the positions, one note is actually sustained while others change. In the following, we focus on vectors without such sustain information. Figure 3 shows that the classical sub-corpus slightly favours chords with less notes (or single notes) than the other genres. The pop genre includes the most frequently chords involving more than 3 strings.

Strings, Frets, and Hand Positions. As expected, the notes on the lowest frets – single notes or notes belonging to a chord – are the most common ones (73% on open strings or on fret ≤ 5 , Figure 4).

Whether they are played alone or within a chord, most pitches are unaltered. On the three lowest strings, all unaltered pitches from E^3 to D^4 are at least twice as frequent as altered notes. More

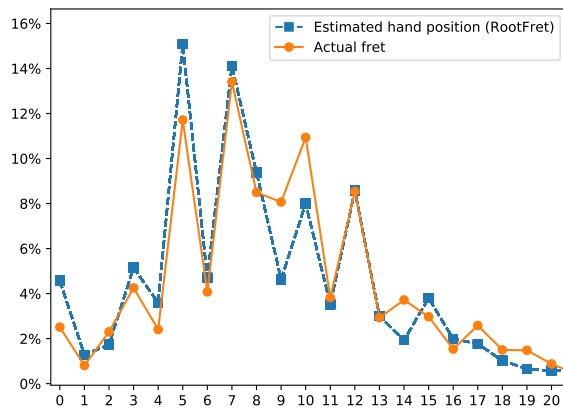


Figure 6: Estimated hand position and actual frets on the sections annotated as *LeadG* in 50 pieces.

generally, the most used pitch classes are unaltered E, G, A, B, C, D (Figure 5). Five of these pitches can be played on open strings (but open strings represent only 17% of the pitches). These pitches belong to the minor pentatonic scales of A and E which are heavily used in guitar tablatures, especially in popular/rock music. They are also present in the keys of A minor, E minor, and in their relative keys G major and C major. On the contrary, the pitch $F\flat$ (in A minor and keys “with flats”) and the pitch $F\sharp$ (in E minor and keys “with sharps”) are almost evenly represented.

Figure 6 compares, for the *MelodicG* section of the corpus, the heuristic estimation of the hand position with the actually played frets. Hand positions on frets 5, 7, 10, and 12 account for about 45% of the positions. Positions 5, 7, and 12 lead to transpositions by a fourth, fifth, and octave compared to the top of the fretboard and hence imply adding at most one alteration compared to the same patterns played at the top of the fretboard. Further research could investigate whether some frets tend to be played from lower positions – such as perhaps frets 9 or 10 from position 7 – thus using fingers rather than shifting the whole hand along the neck.

4.2 Single Notes

Single Notes and Lead/Rhythm Guitar. Single notes account for 47.6% of the vectors. Figure 3 further shows that, as expected, *LeadG* sections include almost single notes (> 90%). *RhythmG* sections however, include a smaller but significant number of single notes (43%), notably in broken chords or on bass notes that are strummed in isolation.

Intervals on Same and Adjacent Strings. Figure 7 shows which (relative) frets are used when one plays two successive notes on the same or on two adjacent strings. On same strings, beyond repetitions of the same fret, that account for 46% of these notes, shifts -2 and $+2$ (major seconds) are favored. On adjacent strings, this confirms that guitarists perform unusual gestures around the G/B pair (compared to other adjacent pairs of strings), due to the interval of major third (compared to the perfect fourth between other adjacent pairs of strings). For example, when going from one note on the G string to another note on the B string, the shift

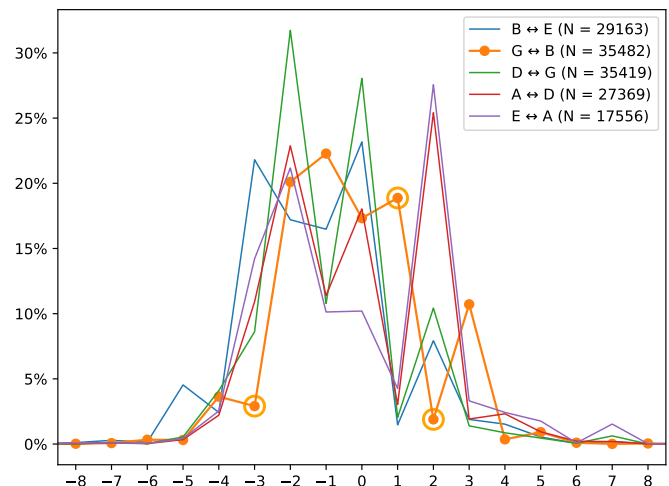
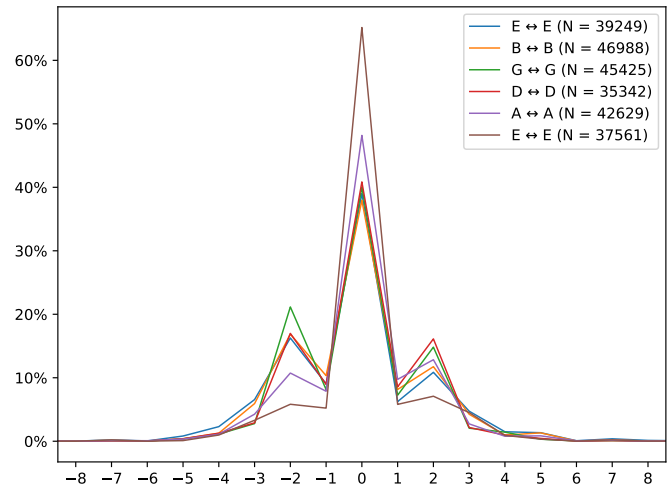


Figure 7: Fret shifting while playing two successive single notes on the same string (top) or on two adjacent strings (bottom). For each string or pair of strings, N is the number of such successions. The circled values highlight shifts on adjacent strings where there is more than a 15% difference between their ratio and the average of the ratios of the other strings: Only shifts between the adjacent strings G^4 and B^4 are concerned.

$+2$ (augmented fourth) is rare while the shift $+3$ (perfect fifth) is favored, contrarily to the other pairs of strings ($+2$: perfect fifth, favored, and $+3$: augmented fifth, more rare).

4.3 Chords: Absolute and Relative Positions

2- and 3-note Chords. Between chords with two or three strings, the most frequent chord is the *power chord* “5” with only a fifth (such as E A), or a fifth then a fourth (E A E). The *fourth chord* (“4”, only a fourth and no fifth nor other intervals) is also very frequent (Figure 8, top). More generally, chords without a third, weakening or preventing a “tonal feel”, are very common in the metal (82 % of the

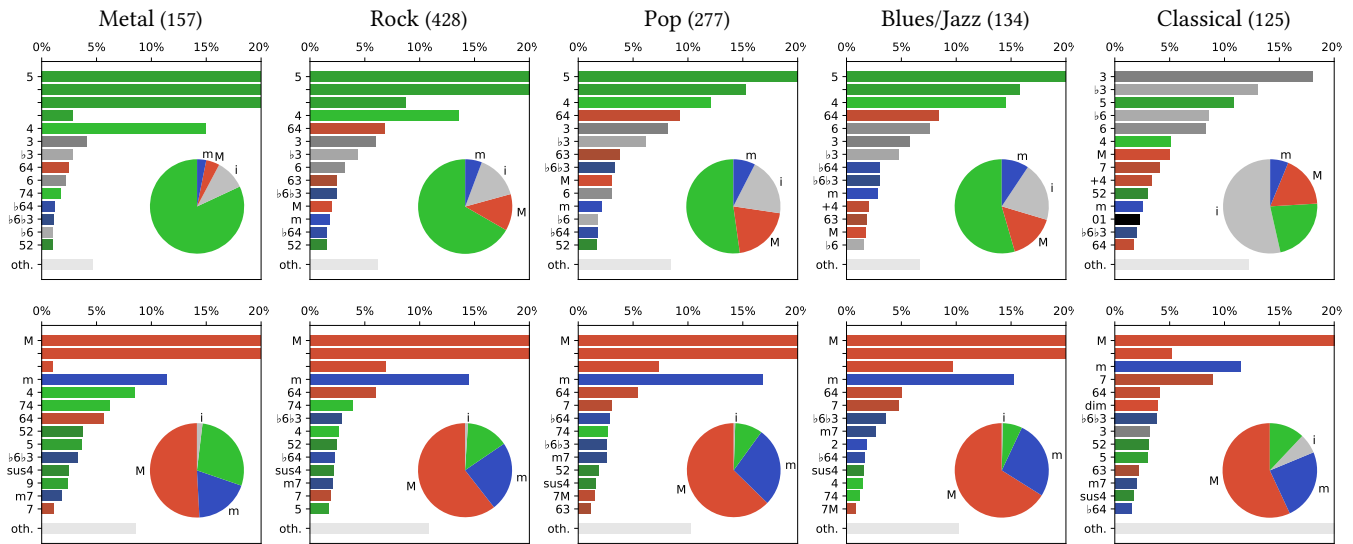


Figure 8: Frequencies of chords in several genres, considering chords with 2 or 3 strings (top), or with 4, 5, or 6 strings (bottom), and gathered into families: major, minor, thirds and sixths, chords without thirds. After each genre is displayed the number of pieces in that genre. To allow to compare more precisely less frequent chords, data over 20% concerning the most frequent chord are split on several lines.

■ major (M)
■ minor (m)
■ thirds/sixths (i)
■ without thirds

2- or 3-note chords) and rock (67%) genres, somewhat less common in the pop (52%) and blue/jazz (54%), and far less common in the classical genre (22%). Limiting the chord to such components can be explained in by the use of *distortion* effects that add a lot of partials to the sound [13], especially used when metal/rock guitarists plays power or fourth chords at lower positions.

On the contrary, reflecting classical harmony, 2- or 3-notes chords in the classical genre are most of the time thirds or sixths, possibly doubled (53% of these chords, compared to between 10% and 20% in the other genres).

4-, 5-, and 6-note Chords. Chords played on four strings at least account for 19.4% of the vectors, and 56% of these chords have at least one open string (data not shown).

Major chords are the most common of these chords (Figure 8, bottom). The diminished seventh chord is noteworthy, accounting for 4.9% of the “classical” genre (including romantic pieces), 0.4% in the jazz/blues, and less than 0.1% in the other genres.

Comparing the most frequent vectors in *RelativeStringFret* and *RelativeStringFret** illustrates the quasi-isomorphism of the guitar. The most frequent chord position between chords with at least four strings, unexpectedly, is a *strict* subset of the standard position for A major (Figure 9), without the highest sounding open string E^5 . We call $[A^*]$ this “incomplete” position (9.05% of such vectors) – note that the transcribers of *MySongBook* usually favor the accuracy of the voicing and fingering over the usage of standard or easier chords.

The complete standard position $[A]$ appears only at rank seven (2.83%). We see three possible reasons for favoring the incomplete position:

- This position still includes a full major triad, with a major third in the upper voice (for the A^* chord, a $C\#^5$ on the B string). This third interval between the lowest and the highest sounding string gives a better voicing than with the fifth E^5 of the complete position;
- In the rhythm accompaniment parts, chords are sometimes played omitting the highest string, in order to maintain accompaniment in a register distinct from that of the melody;
- Some guitarists are accustomed to fret the three strings located on the same fret with the ring finger alone. This makes this position more accessible, especially when playing *standing up*, but makes it harder to play the highest string (Figure 10).

As expected, many of the most frequent positions can be easily transposed on the fretboard (Figure 9). This is the case of the relative position $[E]$ which groups together 7.52% of these vectors, including the open chord E (2.34%), the barre chord F (1.84%) as well as the other barre chords on that position. This is also the case for $[A]$, $[Am]$, and $[Em]$, and for related incomplete positions. On the contrary, other open chords are more difficult to transpose as barre chords, as the $[D]$ which groups 4.94% of the vectors, almost all of which (4.91%) being the open D chord. A similar phenomenon is observed for the relative positions $[C]$ and $[G]$.

						+3		+1				+3	+5	+7	+2		
E	—2—	—0—	—0—	—3—	—0—	—	—	—0—	—0—	—3—	—	—0—	—1—	—0—	—	—0—	
B	—3—	—1—	—0—	—3—	—1—	—2—	—2—	—2—	—0—	—0—	—1—	—0—	—3—	—0—	—2—	—2—	—1—
G	—2—	—0—	—1—	—0—	—2—	—2—	—2—	—2—	—1—	—0—	—0—	—0—	—2—	—1—	—2—	—2—	—2—
D	—0—	—2—	—2—	—0—	—2—	—2—	—2—	—2—	—2—	—0—	—2—	—2—	—0—	—2—	—2—	—2—	—2—
A	—	—3—	—2—	—2—	—0—	—0—	—0—	—0—	—2—	—2—	—3—	—2—	—	—2—	—0—	—0—	—0—
E	—	—	—0—	—3—	—	—	—	—	—0—	—3—	—	—0—	—	—0—	—	—	—
	22906	15005	10904	9733	9424	9287	9192	8909	8572	8545	7285	6805	6446	6344	5867	5067	5051
	4.91%	3.21%	2.34%	2.08%	2.02%	1.99%	1.97%	1.91%	1.84%	1.83%	1.56%	1.46%	1.38%	1.36%	1.26%	1.09%	1.08%
	<u>D</u>	<u>C</u>	<u>E</u>	<u>G</u>	<u>Am</u>	<u>A*</u>	<u>C*</u>	<u>A</u>	<u>F</u>	<u>G</u>	<u>C*</u>	<u>Em</u>	<u>Dm</u>	<u>G</u>	<u>D*</u>	<u>E*</u>	<u>Bm</u>
	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
E	—	—0—	—2—	—0—	—0—	—0—	—0—	—	—	—3—	—3—	—	—0—	—1—	—	—0—	—
B	—2—	—0—	—3—	—1—	—1—	—0—	—2—	—	—1—	—3—	—0—	—1—	—0—	—3—	—0—	—0—	—0—
G	—2—	—1—	—2—	—2—	—0—	—0—	—2—	—1—	—2—	—0—	—0—	—0—	—1—	—2—	—1—	—0—	—1—
D	—2—	—2—	—0—	—2—	—2—	—2—	—2—	—2—	—2—	—0—	—0—	—2—	—2—	—0—	—2—	—2—	—2—
A	—0—	—2—	—	—0—	—3—	—2—	—0—	—2—	—0—	—2—	—2—	—3—	—	—	—2—	—	—2—
E	—	—0—	—	—	—	—0—	—	—0—	—	—3—	—3—	—	—	—	—	—	—0—
	42239	35131	23069	22572	15873	14304	13223	10688	10264	9733	8557	7881	7394	6470	4925	4571	4539
	9.05%	7.52%	4.94%	4.83%	3.40%	3.06%	2.83%	2.29%	2.20%	2.08%	1.83%	1.69%	1.58%	1.39%	1.05%	0.98%	0.97%
	[<u>A*</u>]	[<u>E</u>]	[<u>D</u>]	[<u>Am</u>]	[<u>C</u>]	[<u>Em</u>]	[<u>A</u>]	[<u>E*</u>]	[<u>Am*</u>]	[<u>G</u>]	[<u>G</u>]	[<u>C*</u>]	[<u>E*</u>]	[<u>Dm</u>]	[<u>E*</u>]	[<u>Em*</u>]	[<u>E*</u>]

Figure 9: Most frequent positions implying at least four strings encoded with RelativeStringFret (top, absolute positions) and RelativeStringFret* (bottom, relative positions, with omission of the RootFret value), with their number of occurrences and the ratio of their occurrences in the corpus. Positions notated with a star, like A*, are “incomplete”, meaning that at least one string could be added (most of time the top string) to get another usual position. Underlined positions are discussed in Section 4.3.



Figure 10: A C major chord on the third position, where the ring finger alone is on three frets.

5 CONCLUSIONS AND PERSPECTIVES

We studied guitar music through encodings adapted to the transpositional invariance along the fretboard. The RelativeStringFret encoding offers a compressed representation of guitar tablature information, as the same chord positions are used along the fretboard. Statistics using this encoding confirm the quasi-isomorphism of the guitar, as some aspects of the guitar repertoire and its harmonies, notably on the usages of frets and on some chords depending on the genres. Unaltered pitches are favored and the most used relative chord positions can be easily transposed.

This representation of tablatures in such compact binary vectors may be a good encoding to ease machine learning for both analysis and generation tasks. The vectors we distribute for each genre may give additional insights in the analysis, comparison or learning of styles.

Further research could propose more accurate models to estimate the RootFret value to minimize hand movements over the whole piece, and study this corpus according to more functional or stylistic categories. It could be worth to enlarge the corpus, and also to study hand and finger positions when playing with other tunings, or on other corpora for string instruments, fretted or non-fretted. The distributions of notes and chords could also be compared to other instruments to study whether the reported observations are specific to the guitar. Cognitive experiments could finally study at which point guitar players may think according to such moveable patterns.

ACKNOWLEDGMENTS

We thank anonymous reviewers for their insightful comments on an earlier version of this manuscript.

REFERENCES

- [1] Jules Cournot, Louis Bigo, Mathieu Giraud, and Nicolas Martin. 2020. Encodages de tablatures pour l’analyse de musique pour guitare. In *Journées d’Informatique Musicale (JIM 2020)*.
- [2] Michael Scott Cuthbert and Christopher Ariza. 2010. music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data. In *International Society for Music Information Retrieval Conference (ISMIR 2010)*. 637–642. <https://doi.org/10.5281/zenodo.1416114>
- [3] Orchisama Das, Blair Kaneshiro, and Tom Collins. 2018. Analyzing and Classifying Guitarists from Rock Guitar Solo Tablature. In *Sound and Music Computing Conference (SMC 2018)*. 232–238.
- [4] Jonathan De Souza. 2018. Fretboard Transformations. *Journal of Music Theory* 62, 1 (2018), 1–39.
- [5] Nailson dos Santos Cunha, Anand Subramanian, and Dorien Herremans. 2017. Generating guitar solos by integer programming. *Journal of the Operational*

- Research Society* 69, 6 (2017), 971–985. <https://doi.org/10.1080/01605682.2017.1390528>
- [6] Dinko Fabris. 2005. *Performance on Lute, Guitar, and Vihuela: Historical Practice and Modern Interpretation*. Cambridge University Press, Chapter Lute Tablature Instructions in Italy: A Survey of the Regole from 1507 to 1759.
 - [7] Raphael Foulon, Pierre Roy, and François Pachet. 2013. Automatic classification of guitar playing modes. In *International Symposium on Computer Music and Multidisciplinary Research (CMMR 2013)*. 58–71.
 - [8] Anthony Glise. 1997. *Classical guitar pedagogy: A handbook for teachers*. Mel Bay.
 - [9] Gaëtan Hadjeres and François Pachet. 2016. DeepBach: a Steerable Model for Bach chorales generation. *CoRR* abs/1612.01010 (2016). arXiv:1612.01010 <http://arxiv.org/abs/1612.01010>
 - [10] Gen Hori, Hirokazu Hameoka, and Shigeki Sagayama. 2013. Input-Output HMM Applied to Automatic Arrangement for Guitars. *Journal of Information Processing* 8, 2 (2013), 477–484.
 - [11] Christian Kehling, Jakob Abeßer, Christian Dittmar, and Gerald Schuller. 2014. Automatic Tablature Transcription of Electric Guitar Recordings by Estimation of Score-and Instrument-Related Parameters. In *International Conference on Digital Audio Effects (DAFx 2014)*. 219–226.
 - [12] Otso Lähdeoja, Loïc Reboursière, Thomas Drugman, Stéphane Dupont, Cécile Picard-Limpens, and Nicolas Riche. 2012. Détection des Techniques de Jeu de la Guitare. In *Journées d'Informatique Musicale (JIM 2012)*. 47–53.
 - [13] Esa Lilja. 2015. Dealing with the 3rd: Anatomy of distorted chords and subsequent compositional features of classic heavy metal. In *Modern heavy metal – Markets, practices and cultures*. Aalto University Helsinki, 393–403.
 - [14] Matt McVicar, Satoru Fukayama, and Masataka Goto. 2014. AutoLeadGuitar: Automatic Generation of Guitar Solo Phrases in the Tablature Space. In *International Conference on Signal Processing (ICSP 2014)*. 599–604.
 - [15] Matt McVicar, Satoru Fukayama, and Masataka Goto. 2014. AutoRhythmGuitar: Computer-aided composition for rhythm guitar in the tab space. In *International Computer Music Conference (ICMC 2014)*. 293–300.
 - [16] Kumar Ashis Pati and Alexander Lerch. 2017. A Dataset and Method for Guitar Solo Detection in Rock Music. In *International Conference on Semantic Audio (AES 2017)*. P2–3.
 - [17] Ricardo Scholz and Geber Ramalho. 2008. COCHONUT: Recognizing Complex Chords from MIDI Guitar Sequences. In *International Conference on Music Information Retrieval (ISMIR 2008)*. 27–32.
 - [18] Nicholas Jordan Shea. 2020. *Ecological Models of Musical Structure in Pop-rock, 1950–2019*. Ph.D. Dissertation. The Ohio State University.
 - [19] Michael Stein, Jakob Abeßer, Christian Dittmar, and Gerald Schuller. 2010. Automatic detection of audio effects in guitar and bass recordings. In *Audio Engineering Society Convention 128 (AES 2010)*.
 - [20] Li Su, Li-Fan Yu, and Yi-Hsuan Yang. 2014. Sparse Cepstral, Phase Codes for Guitar Playing Technique Classification. In *International Society for Music Information Retrieval Conference (ISMIR 2014)*. 9–14.
 - [21] Daniel Tompkins. 2015. FRETworks: Idiomatic Voice Leading on the Guitar. In *SMT annual conference in St. Louis*.
 - [22] Music Trades (Ed.). 2021. *The Music Industry Census*.
 - [23] Daniel R. Tuohy and W.D. Potter. 2006. Guitar Tablature Creation with Neural Networks and Distributed Genetic Search. In *Advances in applied artificial intelligence*.
 - [24] Erich M. Von Hornbostel and Curt Sachs. 1914. Systematik der Musikinstrumente. Ein Versuch. *Zeitschrift für Ethnologie* 46, H. 4/5 (1914), 553–590.
 - [25] Andrew Wiggins and Youngmoo Kim. 2019. Guitar tablature estimation with a convolutional neural network. In *International Society for Music Information Retrieval Conference (ISMIR 2019)*. 284–291.
 - [26] Howard Wright. 1995. The Guide To Tab Notation. <https://www.classtab.org/tabling.htm>.
 - [27] Qingyang Xi, Rachel M Bittner, Johan Pauwels, Xuzhou Ye, and Juan Pablo Bello. 2018. GuitarSet: A Dataset for Guitar Transcription.. In *International Society for Music Information Retrieval Conference (ISMIR 2018)*. 453–460.
 - [28] Gary Yim. 2011. *Affordant Chord Transitions in Selected Guitar-Driven Popular Music*. Ph.D. Dissertation. The Ohio State University.