Musical Statistics on the 20er Organ Scale

Harald M. Müller

A musical-statistical question:

On an evening around New Year, after having arranged my 120th or so arrangement for 20er music, a question suddenly occurred to me: How many of all the possible 20er chords are actually used in 20er music? The question popped into my mind because the arrangement I had written was not particularly interesting—it was a "contract arrangement" (what I do for money, but not always with the utmost enthusiasm, I must confess); and I had heard quite a lot of similar arrangements of various "hits" over the years. Somehow, I suddenly felt guilty that I did not use "all of the possibilities of the 20er"—but then, is this possible at all? Certainly when you think of the 88-note piano keyboard or the extensive scales of orchestral instruments, there is no chance that all possibilities of those instruments will ever be covered. On the other hand, the tiny 20er scale … I scratched my head and thought a little more.

To make the question more concrete, and to limit the possibilities even more, I limited myself in two ways:

Let us assume that our 20er organ has no stop that can be drawn—in other words, there is only one sound for each note that can be played.

Moreover, let us only consider chords—i.e. notes played at the same time: The possible sequences of notes are ignored.

Then, I could ask a quite concrete question: *How many of the different possible chords on a single-stop 20er are used in 20er arrangements?*

How many different chords are there?

Before I delve into the question how many chords are used, it is of course necessary to know how many possible chords there are at all. Well, how many? Would you like to guess? Ah, but we have mathematics to help us with this!

- How many single notes are there? 20, of course (in 20er music)!
- How many two-note chords (or "intervals")? We can find this out as follows: We take a first note: There are 20 possibilities for this. Then, we take a second note from the remaining 19. So we end up with 20 times 19 or 380 possibilities. But wait—there's something wrong here: If we take e.g. first c' and then d", we will get the same interval as when we take d" first, and then c'. So we have counted each interval twice—we have to halve 380, and hence get 190 different intervals.
- How many 3-note chords ("triads")? Again we can follow the approach above: First, take one of 20; then, one of the remaining 19; then, one of the remaining 18. This yields 20 times 19 times 18 = 6840 possibilities. But again, we will select each possible triad multiple times—here, there are six different possibilities for each one. Take e.g. c'e'g': We will have selected it as c'e'g' or c'g'e'; e'c'g or e'g'c'; g'c'e' or g'e'c'. The overall number of triads is therefore 6840 divided by six or 1140.
- How many 4-note chords? It's 20 · 19 · 18 · 17 divided by 1 · 2 · 3 · 4 or 116280 divided by 24, which yields 4845.
- How many 5-note chords? 20 · 19 · 18 · 17 · 16 / (1 · 2 · 3 · 4 · 5) = 15504
- 6-note chords: 38760
- 7-note chords: 77520
- 8-note chords: 125970
- 9-note chords: 167960
- 10-note chords: 184756
- 11-note chords: The number of possibilities to select the 11 notes occurring in a chord from a set of 20 notes is just the same as selecting the 9 notes not occurring. So there are just as many 11-note chords as there are 9-note chords: 167960. The same argument also holds for the remaining chord sizes: number of 12-note chords = number of 8-note chords = 125970, number of 13-note chords = number of 7-note chords = 77520 etc.
- There are 20 different 19-note chords (can you write them down?).
- Finally, there is a single, horrible, air-demanding, crashing 20-note chord!

Ahhh … So much for the math!

Poking around a little!

Before we dive into the actual full analysis of that question above, let's poke around a little in some music and look with what we come up—this is more fun than going straight to the difficult parts (it won't be that "difficult"); and maybe also wettens your lips for the real "challenges" to come. So let us think of some chords that might or might not occur in 20er music. (As a pre-requisite, we must agree how to notate our examples. Customarily, 20er organs are tuned to the key of B major. Thus, the scale is as seen in **Figure 1**. All the examples in this text will be written using this scale.

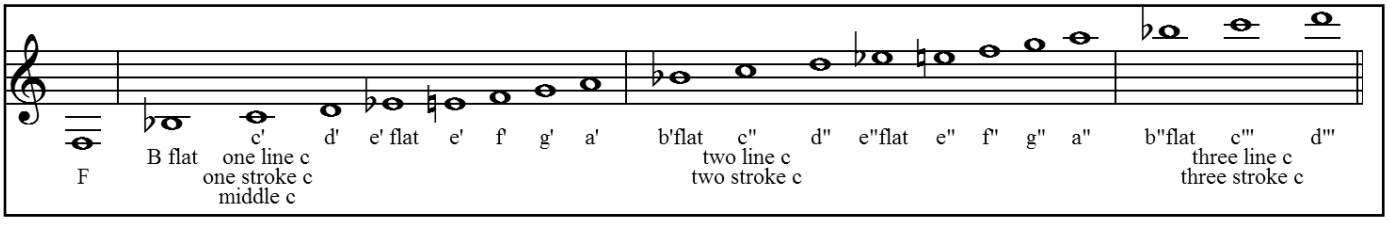
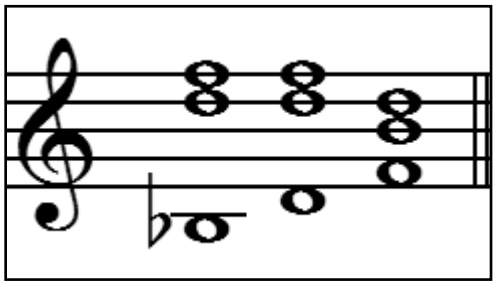
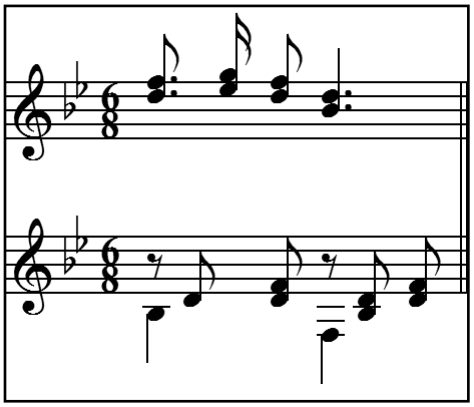


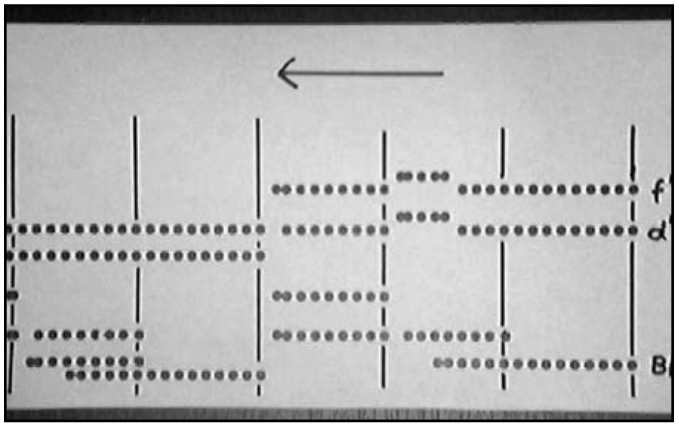
Figure 1. The scale for 20er organs (key of B major).



As a starter, we select some simple ones from the b-flat scale, (**Figure 2**, left).

Here is the beginning of my arrangement of *Silent Night*—first, as a score (**Figure 3**, right**).**





Above, in **Figure 4,** you can see the beginning of the same arrangement as it is punched on a roll; the b-flat chords are marked by black lines.

But of course, you knew that these chords would appear in 20er music—not only here, but at thousands of other places. So let's be a little more adventurous: What about the following chords in **Figure 5?** I started looking for them in my arrangements, but after a short time I gave up—this is too tedious a job! I needed a different method, an automatic method of looking for chords!

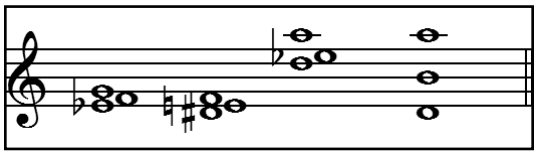


Figure 5. Four strange chords.

How to Count Chords:

Computer to the rescue! All my 120 or so arrangements for 20er scale are written as MIDI files[1]—so computer programs can easily read them and then "do something" with them. Instead of writing a new program to count the chords, I used the punching-machine-controlling program I had written for a friend of mine. That program, called *MusicPunchingInJava*, can be equipped with different "drivers" for different punching machines, which came in handy in my case. I wrote a small driver which emits the notes to a standard computer text file (instead of sending commands to a real punching machine). Here is a small and somewhat simplified sample output of this program—this is again the beginning of *Silent Night* we have seen above:

```
--c:\projects\private\mymusic\19991214-
20er.StilleNacht.mid
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 03    >          d  f
    00:00 04    >  D       d  f
    00:00 04    >  D       d  f
    00:00 02    >  D
    00:00 04    >  D       s  g
    00:00 03       D       s  g
    00:00 02               s  g
    00:01 00
    00:01 04       D  F    d  f
    00:01 04       D  F    d  f
```

The notes are coded as simple letters (or sometimes symbols like > for the deep b-flat, when I ran out of letters; I use "s" for e-flat, because in German, this note is called "es"). On the left, you can see the minutes and seconds from the start of the arrangement—this will help later when a chord has to be identified in a score; and additionally the number of notes in each chord.

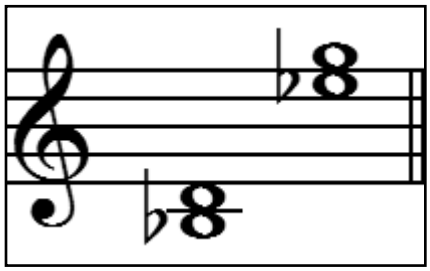
When I actually ran the "punching program" over the 120+ arrangements, this yielded roughly a million lines of chords[2] . A few additional tiny programs helped to me to compress this load of information into a small file, which contained all the existing chords only once. For the fragment of *Silent Nigh*t above, this would look about as follows:

```
00
02      >D
02      sg
03      >df
03      Dsg
04      >Ddf
04      >Dsg
04       DFdf
```

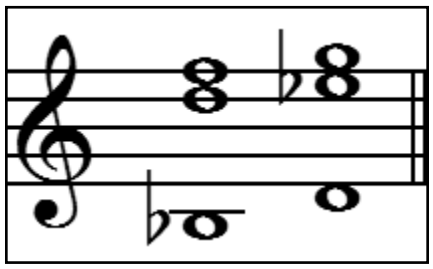
Here you can see that the fragment contains

•the "empty chord" (no note sounds)

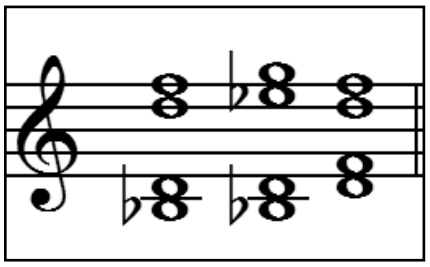
•two 2-note chords ("intervals"), namely (**Figure 6**, right);

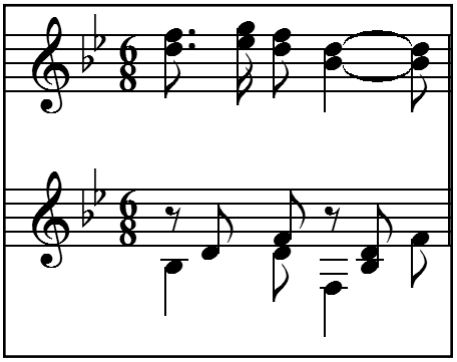


•2 different triads (**Figure 7**, right);



•Finally, 3 different 4-note chords (**Figure 8**, right).

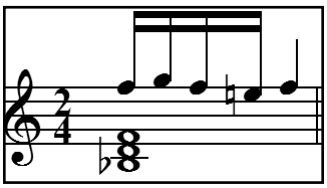




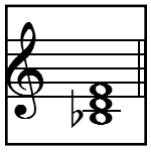
This means for example that the second and third of the b-flat chords which we looked for in the first note example are missing from the following fragment—but how can that be? After all, all three chords are plainly visible in the score (**Figure 9**, above).

"A chord occurs" - what does that mean?

Well, there is a slight problem here with the term "a chord occurs (in an arrangement)." In the following piece of music,



(**Figure 10**, left) everyone would agree that the chord (**Figure 11**, below right) does occur. The fact that there are additional notes sounding, e.g. melody or counter-melody, does not invalidate this observation. However, the trivial programs I had written only checked the occurrence of complete chords, i.e., they would miss a chord if other notes sneaked in between or above or below the chord notes!



What I actually wanted to count, however, were the "really occurring chords," where "really occurring" means the following: a chord *occurs really* in a piece of music if its notes sound at the same time—but possibly together with other notes!

How to Really Count Chords:

A little additional program helped to do the real count. In that program, I first split each chord in each arrangement into all possible sub-chords, as follows: When the file contains e.g.

04 DFdf

(the last line from the *Silent Night* output above), the program would transform this into the following list of sub-chords:

01 D
01 F
01 d
01 f
02 DF
02 Dd
02 Df
02 Fd
02 Ff
02 df
03 DFd
03 DFf
03 Ddf
03 Fdf
04 DFdf[3]

As you can see, these are all the possible sub-chords of the chord DFdf. Now, I could let the program run to find all sub-chords of all chords of **all** the arrangements. A small extension of the program finally compared all the sub-chords to a list of all possible chords of each size.

First results!

Here is the result of all that work:

- Each note occurs somewhere (would you have believed that?!)
- Also, each interval (or 2-note-chord; or pair) does occur somewhere.
- Of all the 1140 possible 3-note chords (triads), only 39 do not occur[5]. Now, this is rather astonishing: Only a mere 3.33% of the three-note possibilities are missing. In other words, the chances that I'd find the strange triads you saw above somewhere in my arrangements was quite high!+I'd have to look for them! But first, let's continue a little with the statistics:
- Of all the 4845 possible 4-note chords, 1307 do not occur—still, I had missed only around 28% of all possible 4-noters.

From here on, the percentages of unused chords rise rapidly:

- Around 10933 or 72% of the 15504 5-noters are missing.
- Only 2529 of the 38760 possible 6-noters are used—more than 93% never occur in my arrangements.
- 7-note-chords: 827 of 77520 possible used—more than

98.5% unused.
- 8-note-chords: 186 of 125970 possible used—more than 99.85% unused.
- 9-note-chords: 27 of 167960 possible used—more than 99.98% unused.
- 10-note-chords: Only two of the possible 184756 are used at all (99.999% unused), as seen in **Figure 12**, below.

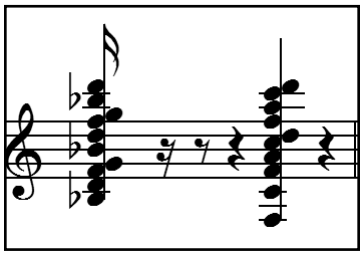
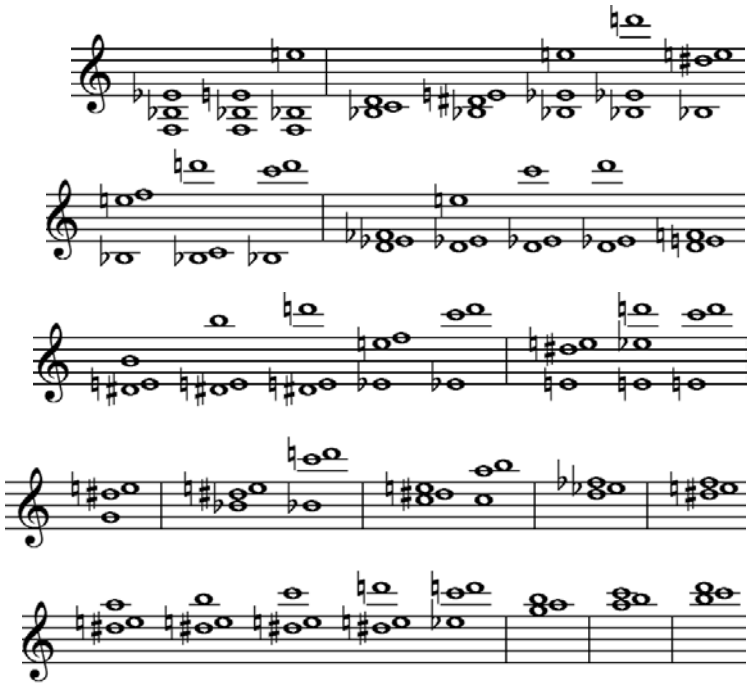


Figure 12.

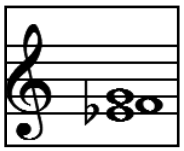
Both occur as short final "crash chords," the first in a medley of film music, the second in an arrangement - which I wrote a short time ago—of a folk song from Berlin, Germany called *Komm Karline*.
- And no 11-noters or larger chords are used anywhere.

The Missing Chords

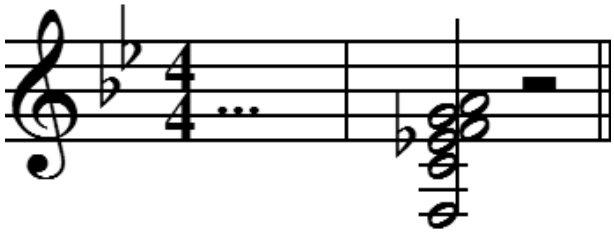
What are the missing chords? - certainly I wanted to know which chords I had missed in the past! For the sake of brevity, I'll only list the 39 missing triads. As a challenge, you (and I) can try to use exactly these chords in future arrangements! When writing the 39 chords, I had sometimes to write d-sharp instead of e-flat, and sometimes even f-flat instead of e-natural (**Figure 13**, below).



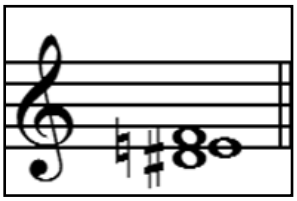
Occurrences of Strange Triads—and a (Small) Disappointment:



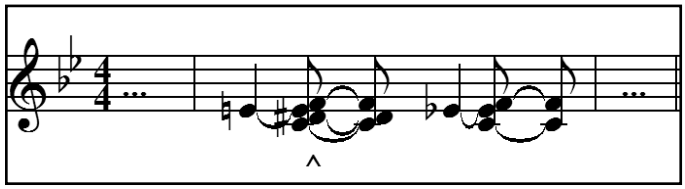
But what about chords that do occur? Let's look for the four "strange chords" you saw at the beginning with the first in **Figure 14**, right).



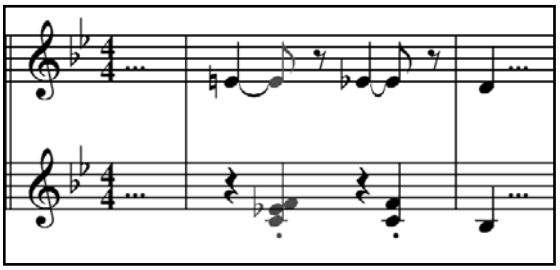
This chord (**Figure 15**, above) is actually part of the long final chord of *Honky Tonk Train* by Meade Lux Lewis.



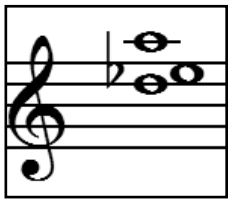
This second strange chord (**Figure 16**, above) occurs in *Grandpa's Spells* by Jelly Roll Morton, measure 73



Now this line of music looks outright chaotic (**Figure 17**, above)—but when notated like what is seen in **Figure 18**, below, it suddenly makes much more sense!



The third strange chord is one I used in a Jazz variation of a German-French chanson called *Quand on n'a pas ce qu'on aime* (**Figure 19 and 20**, right and next page). . .



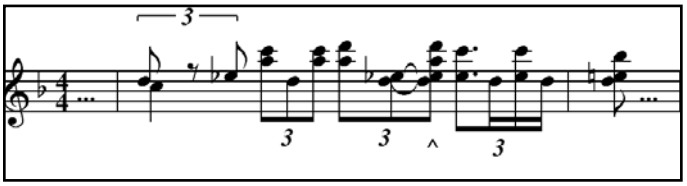
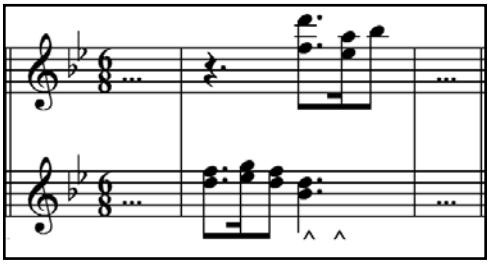
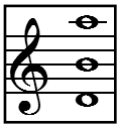


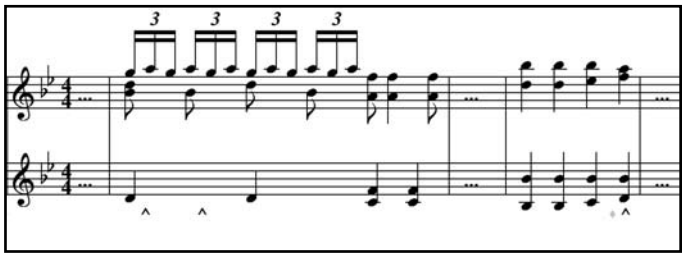
Figure 20.

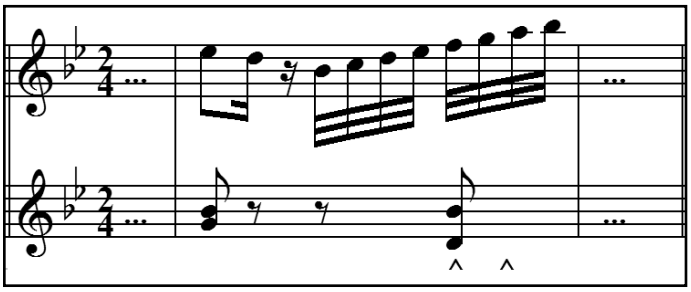
But it also occurs in *Silent Night*, seen in **Figure 21** (below) which made me scratch my head a little: Is this really an occurrence of the chord d+e-flat+a? Mhm …





Let us look at the last chord (**Figure 22**, left) of the four strange ones: there are three places where this chord occurs—or does it? The first two are taken from Scott Joplin's *Maple Leaf Rag* (**Figure 23**, below), the last one is from F.J.Haydn's *Flötenuhr No.2* from 1792 (**Figure 24**, below):





Figures 23 and 24.

I think you will agree that the first and third snippet are not really examples of a occurrences of that chord: Rather, they are coincidences where, during a scale or a trill, the respective notes sound at the same time for a very short time. But no-one would hear this as an occurrence of a chord!

Scales and Trills!

I set out to remove such "wrong occurrences" from my statistics. For this, I again had to rewrite part of my programs (in this case, the "driver" of the punching program). I used the following heuristics:

- If a note is shorter than about 1/6th of a second; and there is a note immediately before it one semitone or one full-tone higher; and there is a note immediately after it one semitone or full-tone lower: Then the note is in a scale!
- The same is true if the note before is lower; and the note after is higher.
- If both the note before and the note after are lower, the note is a trill note (but if both are higher, it is not: The lower note in a trill is usually heard as a kind of "base note").

I call such occurrences of notes "marginal occurrences." Below (**Figure 25**) are examples of the three possible marginal occurrences of notes:

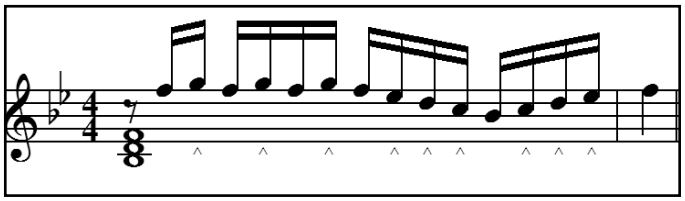


Figure 25.

Of course, one can refine or change these definitions. Actually, they are a compromise between my musical understanding and the possibility to define them so that a program can find them in the scores.

One completely different possibility would be to define that a chord only occurs if all its notes started at the same time. However, this would mean that in the following snippets (**Figure 26**), there occurs no complete C major triad—which is not really satisfying, in my opinion:

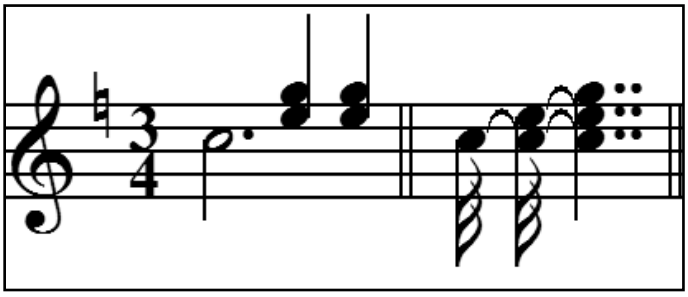


Figure 26.

Still, it might be interesting to find out how many such "synchronous occurrences of chords" are used in 20er music.

Improved results!

Anyway, here are the refined statistics, this time as a simple table:

Chords' size	*Missing chords*	*Percent not used*
1	0 of 20	0%
2	0 of 190	0%
3	66 of 1140; used 1074	5.8%
4	1739 of 4845; used 3106	35.9%
5	11944 of 15504; used 3560	77.0%
6	36859 of 38760; used 1901	95.1%
7	76884 of 77520; used 636	99.2%
8	125819 of 125970; used 151	99.88%
9	167936 of 167960; used: 24	99.98%
10	184754 of 184756; used 2	99.999%
11 and above		100.00%

Single notes and intervals still are used fully. But there are now 66 not-used triads, or 27 more than before. I manually searched for all these 27 triads and found that 16 of them actually occur at a "chord-like" place—so my program is too restrictive when tagging a note occurrence as marginal. Here, for example, is measure 26 from *Alte Kameraden* (**Figure 27**). My program thought that the e-natural was a marginal occurrence, namely a scale occurrence in the "scale" f to e natural to e-flat. However, I overrule this because for me, this is not a marginal note, but a real one—you actually hear the rasping e-natural+e-flat:
(Score 24 ScaleButNotReally)

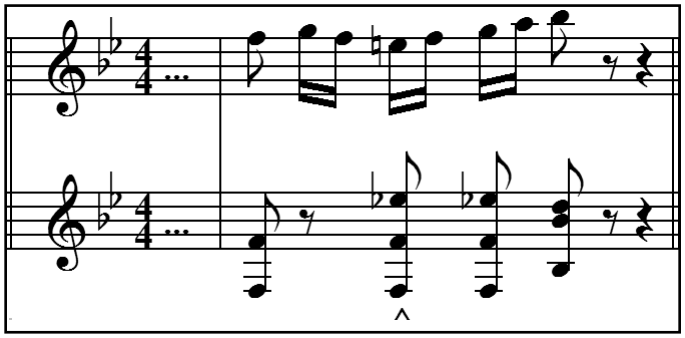


Figure 27. Measure 26 from *Alte Kameraden.*

So actually, the third line of the table above should be:

3	50 of 1140; used 1090	4.4%

I did not do the same tedious checking for the 1739 - 1307 = 432 four-noters my program now tags as marginal occurrences, much less for the five-noters and above. But I am sure there are many more chords that I would like to tag as really occurring, so the percentages of not used chords as shown above are certainly somewhat too high. Looking for this, however, would certainly be a tedious job—and because this would really stretch my, and most probably also your, patience, this is the end of my ponderings about 20er music and statistics!

My conclusion?

It might not be the most important question in music I have tried to shed some light on—but it was definitely fun to find out more about the "statistics of 20er music." And it was, at least for me, astonishing that I have used almost all of the possible three-note chords in my arrangements. Being a perfectionist (well, somewhat …), I'll definitely find the right spots where I can sneak in the remaining 37 or 50 chords so that I can claim a "100 percent triad score!" And, as a long-term goal, why not aim for using the missing 1307 (or 1739, depending on the statistics one chooses) four-note chords also in some arrangements? I'll certainly write a few more of those in the years to come!

The basis for this article was an original inclusion in Mechanical Music Digest, January 5, 2005, entitled *Musico-Statistics on the 20er.*

Notes

1. Actually, I use the *Noteworthy Composer* program to write them. But, like every music score program, Noteworthy can export the files in MIDI format.
2. If you are an alert reader who doesn't believe everything just because it's written somewhere, you will find that there must be something wrong here: in the *Silent Night* example file shown, one (1) second contained 13 chord samples; thus the 1,000,000 lines would have to be 1,000,000/13 = 75,000 seconds long, and each arrangement would have about 75,000/120 = 625 seconds = 10 minutes. That would be awfully long on average for 20er arrangements! The solution of this puzzle: the actual program "sampled" the MIDI files four times as often, so there were about 50 chords per second—which yields about 2 ½ minutes per arrangement, a realistic value.
3. Math question: how many sub-chords are there for a chord with N notes (e.g. N = 4 for DFdf)? I'm sure you can find the formula yourself!
4. If you by chance read the small news to MMD where I first wrote about *Musico-Statistics on the 20er,* you'll have noticed that the numbers are slightly different there. The reason is that I have written a few more arrangements in the meantime, which apparently include a few of the previously missing chords!

Harald M. Müller is an arranger for all sorts of mechanical instruments whose jazz and other arrangements are played around the world. A software engineer by profession and also a hobby mathematician, he tries to combine mathematics, software, and music—but arranging music remains his most prominent interest. Originally from Austria ("where the music comes from . . .") he and his family now live near Munich, Germany.