Mysterium: A Corpus of Alexander Scriabin's Music for Solo Piano

Bryan Jacob Bell
Kent State University
United States of America
bbell28@kent.edu

Abstract

A new digital encoding of 207 works by Alexander Scriabin is reported. The corpus includes all of Scriabin's works for solo piano with an opus number. Each work is in the `**kern` format, having first been encoded into Finale, exported to a MusicXML file, and then converted using the `musicxml2hum` command. The corpus's content and method of encoding are detailed, as well as some complications of the encoding process.

Introduction

Scholars have relied on notated music for centuries to make inferences and judgements about the perception and structure of music. Over the past few decades, fast, powerful computer technologies have emerged, and scholars have used these technologies to their advantage. Contemporary machines provide scholars the opportunity to analyze a large corpus of music quickly, and to apply complicated computational methodologies more efficiently. Computational methods and corpus studies become quite tedious when conducted by hand using traditional Western notation. Empirical and computational musicologists have come to rely on various encoded formats, including the `**kern` representation. In this representation, a musical sequence of notes is encoded in a vertical "spine" of information, in which each new line in the file corresponds to some musical event, such as a note or rest. Notes are represented by lower- and upper-case letters, and rhythms by numbers.

This paper reports the *Mysterium* corpus, a new digital encoding of the solo piano music of Alexander Scriabin (1872–1915). The corpus's content and method of encoding are detailed, as well as some complications of the encoding process.

1 *Mysterium*

The *Mysterium* corpus comprises all of Alexander Scriabin's works for solo piano with opus number and not published posthumously. There are 207 pieces of music in the corpus, with each movement of the multi-movement sonatas (Opp. 6, 19, 23, 30) counted individually. Ninety of the works are preludes, but the corpus also includes sonatas, etudes, mazurkas, impromptus, *poèmes*, and other miscellaneous compositions. By including all of Scriabin's solo piano music published during the composer's lifetime, the *Mysterium* corpus also includes almost all of

Scriabin's music. Only seven major works are absent: the Piano Concerto, Op. 20; the orchestral prelude *Rêverie*, Op. 24; and the five symphonies Opp. 26, 29, 43, 54, and 60. Note that Opp. 50 and 55 are not extant. The corpus's title, *Mysterium*, is taken from an unfinished work by the composer.

Each piece was encoded in the `**kern` representation. `**kern` was chosen because files in this format can be analyzed using the Humdrum Toolkit (Huron, 2002), a set of computer programs that allows for several types of pitch and rhythm analysis. The process for encoding a work was as follows. First, the notes and rhythms were entered into the Finale notation software. Next, the work was exported to a MusicXML file and converted to `**kern` using the `musicxml2hum` command (Sapp, 2017b). Lastly, to check the accuracy of the conversion, the `**kern` file was imported into the Verovio Humdrum Viewer (Sapp, 2017a), an online program that allows one to simultaneously view and edit a `**kern` file and its score notation. Sometimes a segmentation fault error was thrown when running the `musicxml2hum` command. In such instances, the MusicXML file was imported directly into the Verovio Humdrum Viewer, and then converted using the "convert to Humdrum" option. The works were encoded in random order.

```
Error: Inconsistent rhythm analysis occurring near line 653
Expected durationFromStart to be: 1961/12 but found it to be 653/4
Line: 12GG#L 12G#      12rEE    .      12e#L 12b 12dd#
```

Figure 1: Inconsistent rhythm analysis in terminal of Scriabin, Op. 8, No. 12, m. 41.

2 Complications of the Encoding Process

A complication when encoding Scriabin's piano music was that many of the composer's works feature a technique in which temporary secondary and even tertiary voices will enter the musical texture, or that some notes in an arpeggiation, for example, will be prolonged. When either of these scenarios occurs, one must use different "layers" in Finale to have multiple voices on one staff. In a `**kern` file, this means that the spines are often split and re-joined. During the conversion to `**kern` using the `musicxml2hum` command, an "inconsistent rhythm analysis" error similar to that displayed in Figure 1 can happen. In this example from Scriabin's Etude, Op. 8, No. 12, the splitting and re-joining of spines have placed simultaneous events on different lines,

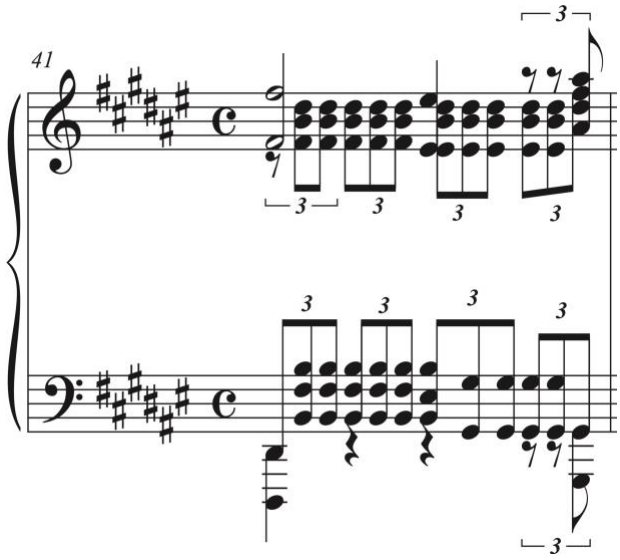


Figure 2: Scriabin, Op. 8, No. 12, m. 41 in traditional notation.

which in turn has created a rhythmic error in which the total rhythmic values for each spine do not equal the value provided by the time signature (see Figure 2 and Figure 3). To correct these errors, the simplest solution was to convert the file, delete the erroneous measures, and then re-encode them manually.

One common feature of Scriabin's music that also creates encoding difficulties is the metric phenomenon of beaming across the bar. While beaming across the bar is supported in Finale, I found that it is often easier to replicate the beaming within a bar as closely as possible in Finale and then manually enter the starting and ending tokens ("L" and "J" respectively, with the number of tokens corresponding to the number of beams) for each beam across the bar into the `**kern` file. Note that the representation of these beams and the accuracy of this representation cannot be verified in Verovio Humdrum Viewer, which currently does not support beaming across the bar.

```
630  12GG#J      12GGG# 12GG#          12f# 12ff#  12bJ 12dd#
631  *v          *v          *         *
632  =41         =41         =41
633  .           2f# 2ff#    12rc
634  *^          *           *
635  12DD#L      4DDD# 4DD#  .         .
636  .           .           .         12f#L 12b 12dd#
637  12BB 12F# 12B           .         .          .
638  .           .           .         12f#J 12b 12dd#
639  12BBJ 12F# 12B          .         .          .
640  .           .           .         12f#L 12b 12dd#
641  12BBL 12F# 12B          4rEE      .          .
642  .           .           .         12f# 12b 12dd#
643  12BB 12F# 12B           .         .          .
644  .           .           .         12f#J 12b 12dd#
645  12BBJ 12F# 12B          .         .          .
646  .           .           4e# 4ee#  12e#L 12b 12dd#
647  12BBL 12E# 12B          4rEE      .          .
648  .           .           .         12e# 12b 12dd#
649  12GG# 12G#  .           .         .
650  .           .           .         8e#J 8b 8dd#
651  12GG#J 12G# .           .         .
652  .           .           12raa     .
653  12GG#L 12G# 12rEE       .         12e#L 12b 12dd#
654  .           .           12raa     .
655  12GG# 12G#  12rEE       .         12e# 12b 12dd#
656  .           .           12a# 12aa#  .
657  12GG#J      12GGG# 12GG#          .          12dd#J 12ff#
658  =42         =42         =42       =42
```

Figure 3: m. 41 of Scriabin, Op. 8, No. 12, displayed in Verovio Humdrum Viewer. At Line 631, the spines have re-joined incorrectly and then, at Line 634, split incorrectly. As a result of this, the onset of the measure has been divided between Line 633 and Line 635, creating rhythmic errors.

3 Implications of the *Mysterium* Corpus for Encoding

Figure 4 displays an excerpt from Scriabin's Etude, Op. 42, No. 8. These two measures (mm. 8–9) are characteristic of multiple metric-rhythmic stylistic trends in the composer's transitional and late periods. Notice first that the onset of each beat does not align with the metric "grid" suggested by the time signature. Rather, the onset of each beat is aligned with the third triplet eighth note of the expected quarter-note beat.

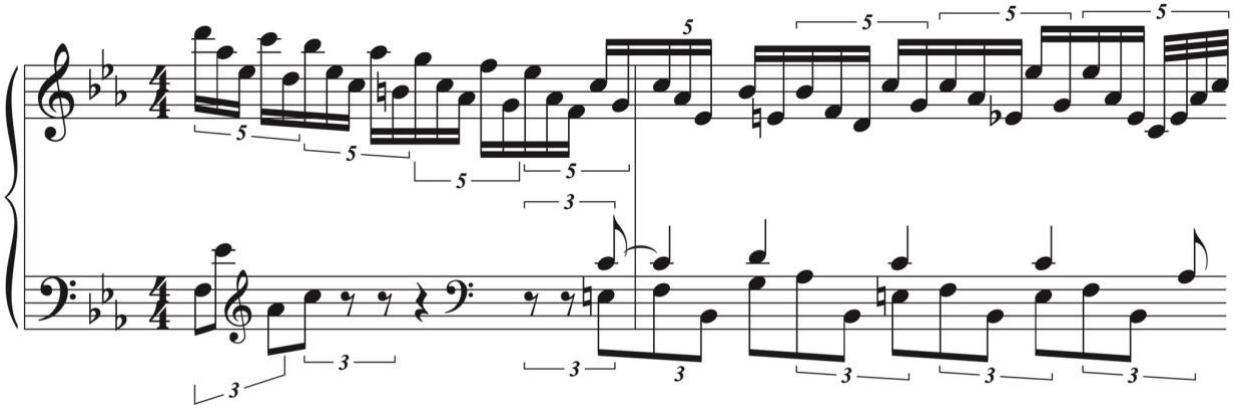


Figure 4: mm. 8–9 of Scriabin, Op. 42, No. 8 (Original notation).

In the `**kern` representation, this complex phenomenon is quite simple to represent; one simply needs to put the onset of the triplet and quintuplet groups on the same line. As for the polyrhythmic alignment between the beats, this is taken care of for the encoder when using a notation software and then converting the file. The encoder could also do this manually by calculating the appropriate number to represent a tuplet note's division of the bar, and by identifying the composite rhythm.

The tenor voice on the lower staff, however, presents an interpretive challenge to the encoder. Because the grouping of each beat leads to beaming across the bar, the tenor's first quarter note—which would otherwise be notated as a normative triplet eighth-note-grouping quarter note—is divided into a triplet eighth note at the end of m. 8 tied over to a quarter note at the beginning of m. 9. The first quarter note of m. 9 then, in Scriabin's original notation, should align with only two triplet eighth notes while the remaining quarter notes are grouped with three triplet eighth notes. And, again, the onset of each quarter note is out of phase with the metric "grid" prescribed by the time signature.

In a notation software like Finale, Scriabin's original notation can be preserved using features such as the ability to put multiple notes in a measure, hide notes and rests, and the note position tool. Such a preservation would only be cosmetic, though—the notes would appear as they do in Scriabin's notation, but the correct onset alignment would not in fact exist. Figure 5 shows an alternate notation in Finale in which the alignment is correct but the notation differs. In the `**kern` representation, Scriabin's notation can be closely preserved as displayed in Figure 6.

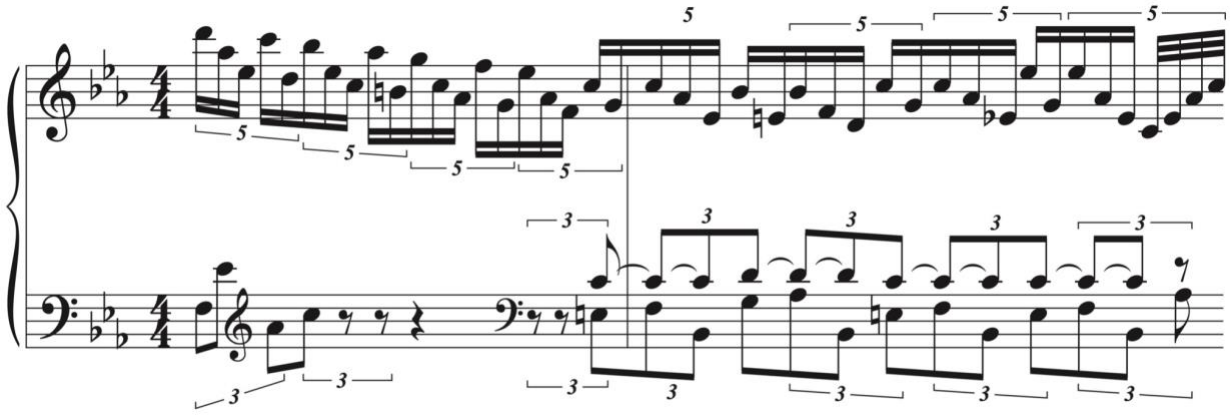
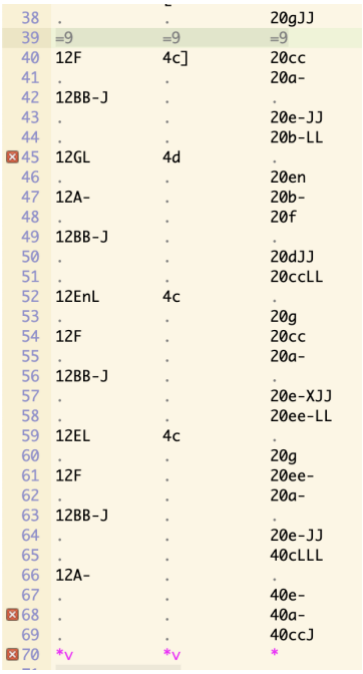


Figure 5: mm. 8–9 of Scriabin, Op. 42, No. 8 (Alternate notation).

Note that Verovio Humdrum Viewer has marked a rhythmic error on Line 45 and would not display the score correctly. This representation should also not be relied upon for a computational metric or rhythmic analysis. The point here is that the traditional score affords more flexibility than a digital representation. A score can be described as *regular*, while a digital representation is *systematic*. When an encoder is confronted with a score prepared by a composer who has manipulated its flexibility, the encoder often is also confronted with a choice to either ensure the representation's musical accuracy—at the possible cost of the composer's notation—or to ensure the representation's notational accuracy at the possible expense of the musical accuracy. This conundrum is worsened when the score is both irregular and ambiguous, because then the implications of the composer's notation for what constitutes "musical accuracy" is unclear.



38	.	.	20gJJ
39	=9	=9	=9
40	12F	4c]	20cc
41	.	.	20a-
42	12BB-J	.	.
43	.	.	20e-JJ
44	.	.	20b-LL
45	12GL	4d	.
46	.	.	20en
47	12A-	.	20b-
48	.	.	20f
49	12BB-J	.	.
50	.	.	20dJJ
51	.	.	20ccLL
52	12EnL	4c	.
53	.	.	20g
54	12F	.	20cc
55	.	.	20a-
56	12BB-J	.	.
57	.	.	20e-XJJ
58	.	.	20ee-LL
59	12EL	4c	.
60	.	.	20g
61	12F	.	20ee-
62	.	.	20a-
63	12BB-J	.	.
64	.	.	20e-JJ
65	.	.	40cLLL
66	12A-	.	.
67	.	.	40e-
68	.	.	40a-
69	.	.	40ccJ
70	*v	*v	*

Figure 6: m. 9 of Scriabin, Op. 42, No. 8 in `**kern` with Scriabin's original notation preserved. Note the red "x" marks on the left side from Verovio Humdrum Viewer, indicating an incorrect rhythm.

Another problem with the regular but not systematic score is consistency. As mentioned earlier, Scriabin's musical textures are often thickened by temporary secondary and tertiary voices or by individual notes in a gesture (e.g., one note in an arpeggiation of four notes) being sustained. In Finale, these events are encoded through multiple "layers" on one staff. In `**kern`, these events are encoded in multiple spines. These layers are not as straightforward as they may be in a fugue, for example, where there is a consistent stratification of one voice (e.g., the soprano) above another voice (e.g., the alto) on a single staff. Rather, in Scriabin's music, there is often a blending of layers achieved by techniques like voice-crossing, voicing notes "in-between" those of another layer, and so on. There is little consistency between each piece, and sometimes between different sections of the same piece; consequently, it is not always clear which voice to enter in a second layer and which to enter in a third layer. When encoding, such inconsistencies in the score lead to inconsistent layer-assignment and inconsistent splitting and re-joining of spines in a `**kern` representation. For the computational analyst, a familiarity with both the composer's original notation and the encoded representation of such a body of music is required *a priori* to any analysis, and one should limit assumptions based on normative notational practices.

Conclusion

This paper has introduced a novel digital encoding of the solo piano music of Alexander Scriabin. Plans for the future of this corpus are twofold. First, more information will be added to the piano music encodings, including dynamic and tempo markings, accents, phrase markings, etc. Second, the seven orchestral works will be encoded to create a full corpus of Scriabin's completed works

published during his lifetime. In discussing some of the difficulties of the encoding process for this type of music—i.e., piano music composed by a Romantic or post-Romantic composer—this paper has addressed some of the issues faced by the encoder and the analyst. For the encoder, the flexibility and ambiguity of the Western score present challenges of accuracy and consistency. Encoding this music does not simply require the encoder to enter the correct information; it requires the encoder to make musical decisions. For the analyst, the use of this corpus requires significant *a priori* study of the scores and the encodings, of their differences and their inconsistencies, and of the potential caveats that they introduce when applying a given methodology or model.

References

Huron, D. (2002). Music information processing using the Humdrum Toolkit: Concepts, examples, and lessons. *Computer Music Journal* 26(2), 11–26.

Sapp, C. (2017a, March 3). *Verovio Humdrum Viewer documentation*. Humdrum. Retrieved April 2, 2021. https://doc.verovio.humdrum.org/.

Sapp, C. (2017b, June 20). *craigsapp/musicxml2hum*. GitHub. Retrieved April 2, 2021. https://github.com/craigsapp/musicxml2hum.