Rekordbox Export Structure Analysis

James Elliott Deep Symmetry, LLC

January 11, 2020

Abstract

The files written to external media by rekordbox for use in player hardware contain a wealth of information that can be used in place of queries to the remotedb server on the players, which is important because they can be obtained from the players' NFS servers, even if there are four players in use sharing the same media. Under those circumstances, remotedb queries are impossible. This article documents what has been learned so far about the files, and how to interpret them.

Contents

l	Data	base Exp	ports
	1.1	File He	ader
	1.2	Table P	Pages
	1.3	Table R	Rows
		1.3.1	Album Rows
		1.3.2	Artist Rows
		1.3.3	Artwork Rows
		1.3.4	Color Rows
		1.3.5	Genre Rows
		1.3.6	Key Rows
		1.3.7	Label Rows
		1.3.8	Playlist Tree Rows
		1.3.9	Playlist Entry Rows
		1.3.10	Track Rows
	1.4	Device	SQL Strings
		1.4.1	Long ASCII Strings
		1.4.2	Long UTF-16 Big-Endian Strings 15
		1.4.3	Short ASCII Strings

CONTENTS 2

2	Anal	ysis Files		16
	2.1	Analysis	s File Header	17
	2.2	-	s File Sections	18
		2.2.1	Beat Grid Tag	18
		2.2.2	Cue List Tag	19
		2.2.3	Extended (nxs2) Cue List Tag	21
		2.2.4	Path Tag	23
		2.2.5	VBR Tag	23
		2.2.6	Waveform Preview Tag	24
		2.2.7	Tiny Waveform Preview Tag	25
		2.2.8	Waveform Detail Tag	25
		2.2.9	Waveform Color Preview Tag	26
		2.2.10	Waveform Color Detail Tag	27
		2.2.11	Song Structure Tag	28
3	Crate	e Digger		31
Lis	st of Fi	gures		31
Lis	st of Ta	ables		32

1 Database Exports

The starting point for finding track metadata from a player is the database export file, which can be found within rekordbox media at the following path:

/PIONEER/rekordbox/export.pdb

(If you are using the Crate Digger FileFetcher to request this file, use that path as the filePath argument, and use a mountPath value of /B/ if you want to read it from the SD slot, or /C/ to obtain it from the USB slot).

The file is a relational database format designed to be efficiently used by very low power devices (there were deployments on 16 bit devices with 32K of RAM). Today you are most likely to encounter it within the Pioneer Professional DJ ecosystem, because it is the format that their rekordbox software uses to write USB and SD media which can be mounted in DJ controllers and used to play and mix music.

The file consists of a series of fixed size pages. The first page contains a file header which defines the page size and the locations of database tables of different types, by the index of their first page. The rest of the pages consist of the data pages for all of the tables identified in the header.

Each table is made up of a series of rows which may be spread across any number of pages. The pages start with a header describing the page and linking to the next page. The rest of the page is used as a heap: rows are scattered around it, and located using an index structure that builds backwards from the end of the page. Each row of a given type has a fixed size structure which links to any variable-sized strings by their offsets within the page.

As changes are made to the table, some records may become unused, and there may be gaps within the heap that are too small to be used by other data. There is a bit map in the row index that identifies which rows are actually present. Rows that are not present must be ignored: they do not contain valid (or even necessarily well-formed) data.

The majority of the work in reverse-engineering this format was performed by Henry Betts¹ and Fabian Lesniak², to whom I am hugely grateful.

1.1 File Header

Unless otherwise stated, all multi-byte numbers in the file are stored in little-endian byte order. Field names used in the Figures match the IDs assigned to them in the Kaitai Struct specification³, unless that is too long to fit, in which case a subscripted abbreviation is used, and the text will mention the actual struct field name.

The first page begins with the file header, shown in Figure 1. The header starts with four zero bytes, followed by a four-byte integer, *len_page* at byte 04, that

https://github.com/henrybetts/Rekordbox-Decoding

²https://github.com/flesniak/python-prodj-link

 $^{^3} https://github.com/Deep-Symmetry/crate-digger/blob/master/src/main/kaitai/rekordbox_pdb.ksy$

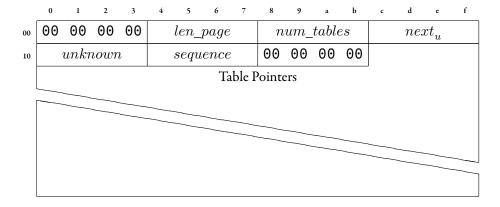


Figure 1: File Header



Figure 2: Table Pointer

establishes the size of each page (including this first one), in bytes. This is followed by another four-byte integer, num_tables at byte 08, which reports the number of different tables that are present in the file. Each table will have a table pointer entry in the "Table pointers" section of the file header, described below, that identifies and locates the table.

The four-byte integer $next_u$ at byte $0\,\mathrm{C}$ has an unknown purpose, but Mr. Lesniak named it next_unused_page and said "Not used as any empty_candidate, points past the end of the file." The four-byte integer sequence, at byte 14, was described "Always incremented by at least one, sometimes by two or three." and I assume this means it reflects a version number that rekordbox updates when synchronizing to the exported media.

Finally, there is another series of four zero bytes, and then the header ends with the list of table pointers which begins at byte 1C. There are as many of these as specified by num_tables , and each has the structure shown in Figure 2.

Each Table Pointer is a series of four four-byte integers. The first, type, identifies the type of table being defined. The known table types are shown in Table 1. The second value, at byte 04 of the Table Pointer, was called $empty_candidate$ by Mr. Lesniak. It may link to a chain of empty pages if the database is ever garbage collected, but this is speculation on my part.

Type	Name	Row Content
0	tracks	Track metadata: title, artist, genre, artwork ID, play-
		ing time, etc., Section 1.3.10.
1	genres	Musical genres, for reference by tracks and search-
		ing, Section 1.3.5.
2	artists	Artists, for reference by tracks and searching, Sec-
3	albums	tion 1.3.2. Albums, for reference by tracks and searching, Sec-
4	labels	tion 1.3.1. Music labels, for reference by tracks and searching,
5	keys	Section 1.3.7. Musical keys, for reference by tracks, searching, and
		key matching, Section 1.3.6.
6	colors	Color labels, for reference by tracks and searching,
7	playlist_tree	Section 1.3.4. Describes the hierarchical tree structure of available
		playlists and folders grouping them, Section 1.3.8.
8	playlist_entries	Links tracks to playlists, in the right order, Sec-
1.2	1	tion 1.3.9.
13	artwork	File paths of album artwork images, Section 1.3.3.
16	columns	Details not yet confirmed.
19	history	Records the tracks played during performances.

Table 1: Table Types

Other than the type, the two important values are $first_page$ at byte 08 and $last_page$ at byte 0C. These tell us how to find the table. They are page indices, where the page containing the file header has index 0, the page with index 1 begins at byte len_page , and so on. In other words, the first page of the table identified by the current Table Pointer can be found within the file starting at the byte $len_page \times first_page$.

The table is a linked list of pages: each page contains the index of the next page after it. However, you need to keep track of the <code>last_page</code> value for the table, because it tells you not to try to follow the next page link once you reach the page with that index. (If you do keep going, you will start reading pages of some different table.) The structure of the table pages themselves are described in the next section.

As far as we know, the remainder of the first page after the table pointers is unused.

1.2 Table Pages

The table header is followed by the table pages themselves. These each have the size specified by len_page in Figure 1, and the structure shown in Figure 3.

Data pages all seem to have the header structure described here, but not all of them actually store data. Some of them are "strange" and we have not yet figured

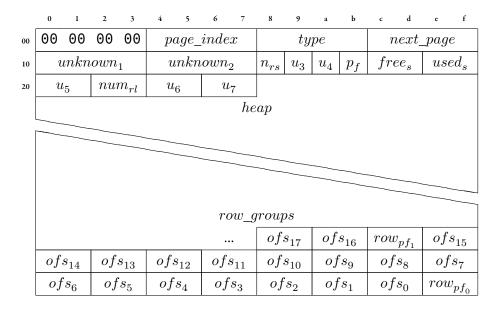


Figure 3: Table Page

out why. The discussion below describes how to recognize a strange page, and avoid trying to read it as a data page.

The first four bytes of a table page always seem to be zero. This is followed by a four-byte value $page_index$ which identifies the index of this page within the list of table pages (the header has index 0, the first actual data page the index 1, and so on). This value seems to be redundant, because it can be calculated by dividing the offset of the start of the page by len_page , but perhaps it serves as a sanity check.

This is followed by another four-byte value, type, which identifies the type of the page, using the values shown in Table 1. This again seems redundant because the table header which was followed to reach this page also identified the table type, but perhaps it is another sanity check, or an alternate way to tell, when following page links, that you have reached the end of the table you are interested in. Speaking of which, the next four-byte value, $next_page$, is that link: it identifies the index at which the next page of this table can be found, as long as we have not already reached the final page of the table, as described in Section 1.1.

The exact meaning of $unknown_1$ is unclear. Mr. Flesinak said "sequence number $(0\rightarrow 1:\ 8\rightarrow 13,\ 1\rightarrow 2:\ 22,\ 2\rightarrow 3:\ 27)$ " but I don't know how to interpret that. Even less is known about $unknown_2$. But num_rows_small at byte 18 within the page (abbrviated n_{rs} in Figure 3) holds the number of rows that are present in the page, unless num_rows_large (below) holds a value that is larger (but not equal to 1fff). This seems like a strange mechanism for dealing with the fact that some tables (like playlist entries) have a lot of very small rows, too many to count with a single byte. But then why not just always use num_rows_large ?

The purpose of the next two bytes are is also unclear. Of u_3 Mr. Flesniak said "a bitmask (first track: 32)", and he described u_4 as often 0, sometimes larger, especially for pages with a high number of rows (e.g. 12 for 101 rows).

Byte 1b is called $page_flags$ (abbrviated p_f in Figure 3). According to Mr. Flesniak, "strange" (non-data) pages will have the value 44 or 64, and other pages have had the values 24 or 34. Crate Digger considers a page to be a data page if $page_flags\&40=0$.

Bytes 1C-1d are called $free_size$ (abbreviated $free_s$ in Figure 3), and store the amount of unused space in the page heap (excluding the row index which is built backwards from the end of the page); $used_size$ at bytes 1C-1d (abbreviated $used_s$) stores the number of bytes that are in use in the page heap.

Bytes 20 - 21, u_5 , are of unclear purpose. Mr. Flesniak labeled them "(0 \rightarrow 1: 2)."

Bytes 20-21, num_rows_large (abbrviated num_{rl} in Figure 3) hold the number of entries in the row index at the end of the page when that value is too large to fit into num_rows_small (as mentioned above), and that situation seems to be indicated when this value is larger than num_rows_small , but not equal to 1fff.

 u_6 at bytes 24-25 seems to have the value 1004 for strange pages, and 0000 for data pages. And Mr. Flesniak describes u_7 at bytes 26-27 as "always 0 except 1 for history pages, num entries for strange pages?"

After these header fields comes the page heap. Rows are allocated within this heap starting at byte 28. Since rows can be different sizes, there needs to be a way to locate them. This takes the form of a row index, which is built from the end of the page backwards, in groups of up to sixteen row pointers along with a bitmask saying which of those rows are still part of the table (they might have been deleted).

The number of row index entries is determined, as described above, by the value of either num_rows_small or num_rows_large .

The bit mask for the first group of up to sixteen rows, labeled row_{pf_0} in Figure 3 (meaning "row presence flags group 0"), is found in the last two bytes of the page. The low order bit of this value will be set if row 0 is really present, the next bit if row 1 is really present, and so on. The two bytes before these flags, labeled ofs_0 , store the offset of the first row in the page. This offset is the number of bytes past the end of the page header at which the row itself can be found. So if row 0 begins at the very beginning of the heap, at byte 28 in the page, ofs_0 would have the value 0000.

As more rows are added to the page, space is allocated for them in the heap, and additional index entries are added at the end of the heap, growing backwards. Once there have been sixteen rows added, all of the bits in row_{pf_0} are accounted for, and when another row is added, before its offset entry ofs_{16} can be added, another row bit-mask entry row_{pf_1} needs to be allocated. And so the row index grows backwards towards the rows that are being added forwards, and once they are too close for a new row to fit, the page is full, and another page gets allocated to the table.

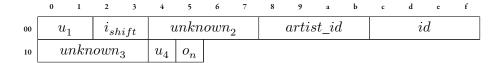


Figure 4: Album Row

1.3 Table Rows

The structure of the rows themselves is determined by the type of the table, using the values shown in Table 1.

1.3.1 Album Rows

Album rows hold an album name and ID along with an artist association, with the structure shown in Figure 4. The unknown value at bytes 00-01 seems to usually have the values 80-00. It is followed by a two-byte value Mr. Flesniak called $index_shift$, although I don't know what that means, and another four bytes of unknown purpose. But at bytes 08-00 we finally find a value we have a use for: $artist_id$ holds the ID of an artist row associated with this track row. This is followed by id, the ID of this track row itself, at bytes 0C-0f. We assume that there are index tables somewhere that would let us locate the page and row index of a record given its table type and ID, but we have not yet found and figured them out.

This is followed by five more bytes with unknown meaning, and the final byte in the row, ofs_name is a pointer to the track name (labeled o_n in Figure 4. To find the location of the name, add ofs_name bytes to the address of the start of the track row itself. The name itself is encoded in a surprisingly baroque way, explained in Section 1.4

1.3.2 Artist Rows

Artist rows hold an Artist name and ID, with the structure shown in Figure 5 or Figure 6. The *subtype* value at bytes 00-01 determines which variant is used. If the artist name was allocated close enough to the row to be reached by a single byte offset, *subtype* has the value 0060, and the row has the structure in Figure 5. If the name is too far away for that, *subtype* has the value 0064 and the row has the structure in Figure 6.

In either case, subtype is followed by the unexplained two-byte value found in many row types that Mr. Flesniak called $index_shift$, and then by id, the ID of this artist row itself, at bytes 04-07, an unknown value at byte 08, and ofs_name_near at byte 09 (labeled o_n in Figure 5), the one-byte name offset used only in the first variant.

If subtype is 0064, the value of ofs_name_near is ignored, and instead the two-byte value ofs_name_far (labeled o_{far} in Figure 6) is used.

0	1	2	3	4	5	6	7	8	9	a	Ь	c	d	e	f
000	60	i_{sh}	ift		i	d		u_1	o_n						

Figure 5: Artist Row with Nearby Name

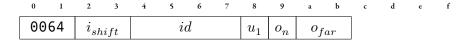


Figure 6: Artist Row with Far Name

Whichever name offset is used, it is a pointer to the artist name. To find the location of the name, add the value of the offset to the address of the start of the artist row itself. This gives the address of a DeviceSQL string holding the name, with the structure explained in Section 1.4

1.3.3 Artwork Rows

Artwork rows hold an id (which tracks refer to) and the path at which the corresponding album art image file can be found, with the structure shown in Figure 7. Note that in this case, the DeviceSQL string path is embedded directly into the row itself, rather than being located elsewhere in the heap through an offset. The structure of the string itself is still as described in Section 1.4.

1.3.4 Color Rows

Color rows hold a numeric color id (which controls the actual color displayed on the player interface) at bytes 05 - 06 and a text label or name starting at byte 08 which is a DeviceSQL string shown in the information panel for tracks that are assigned the color. The rows have the structure shown in Figure 8. There are several

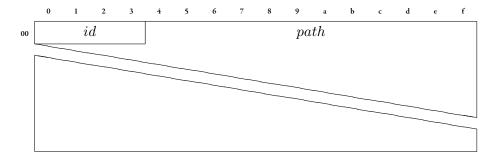


Figure 7: Artwork Row

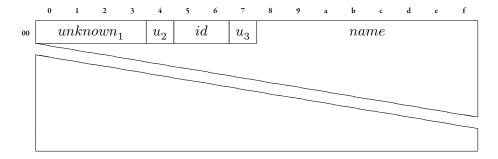


Figure 8: Color Row

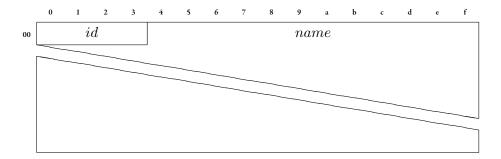


Figure 9: Genre or Label Row

bytes in the row that are not yet known to have any meaning. The structure of name is described in Section 1.4.

1.3.5 Genre Rows

Genre rows hold a numeric genre id (which tracks can be assigned) at bytes 00-03 and a text name starting at byte 04 which is a DeviceSQL string. The rows have the structure shown in Figure 9. The structure of name is described in Section 1.4.

1.3.6 Key Rows

Key rows represent musical keys. They hold a numeric id (which tracks can be assigned) at bytes 00-03 and a text name starting at byte 08 which is a DeviceSQL string. (There seems to be a second copy of the ID at bytes 04-07.) The rows have the structure shown in Figure 10. The structure of name is described in Section 1.4.

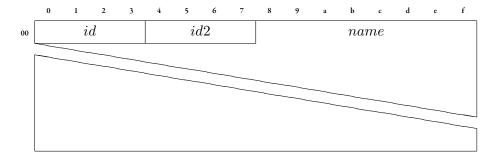


Figure 10: Key Row

1.3.7 Label Rows

Label rows represent record labels. They hold a numeric id (which tracks can be assigned) at bytes 00-03 and a text name starting at byte 04 which is a DeviceSQL string. The rows have the structure shown in Figure 9. The structure of name is described in Section 1.4.

1.3.8 Playlist Tree Rows

Playlist tree rows are used to organize the hierarchical structure of the playlist menu. There is probably an index somewhere that makes it possible to find the right rows directly when loading a playlist, but we have not yet figured out how indices work in DeviceSQL databases, so Crate Digger simply reads all the rows and builds its own in-memory index of the tree.

Playlist tree rows can either represent a playlist "folder" which contains other folders and playlists, or a regular playlist which holds only tracks. The rows are identified by an id at bytes 0C-0f, and also contain a $parent_id$ at bytes 00-03 which is how the hierarchical structure is represented: the contents of a folder are the other rows in this table whose $parent_id$ folder is equal to the id of the folder.

Similarly, the tracks that make up a regular playlist are the Playlist Entry Rows (described in Section 1.3.9) whose $playlist_id$ is equal to this row's id.

Each playlist tree row also has a text name starting at byte 14 which is a DeviceSQL string displayed when navigating the hierarchy, a $sort_order$ indicator at bytes 08-0b (this may be the same value used to select sort orders when requesting menus using the dbserver protocol, shown in Table 7 of the dysentery Protocol Analysis paper, but this has not yet been confirmed), and a value that specifies whether the row defines a folder or a playlist. In the Kaitai Struct, this value is called raw_is_folder , is found at bytes 10-13, and has a non-zero value for folders. For convenience, the struct also defines a derived value, is_folder , which is a boolean.

The rows have the structure shown in Figure 11. The structure of name is

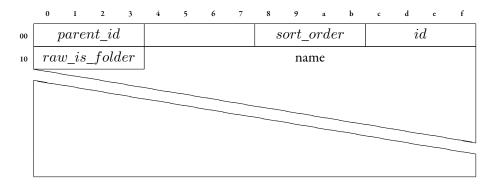


Figure 11: Playlist Tree Row



Figure 12: Playlist Entry Row

described in Section 1.4.

1.3.9 Playlist Entry Rows

Playlist entry rows list the tracks that belong to a particular playlist, and also establish the order in which they should be played. They have a very simple structure, shown in Figure 12, containing only three values. The $entry_index$ at bytes 00-03 specifies the position within the playlist at which this entry belongs. The $track_id$ at bytes 04-07 identifies the track to be played at this position in the playlist, by corresponding to the id of a row in the Track table described in Section 1.3.10, and the $playlist_id$ at bytes 08-0b identifies the playlist to which it belongs, by corresponding to the id of a row in the Playlist Tree described in Section 1.3.8.

1.3.10 Track Rows

Track rows describe audio tracks that can be played from the media export, and provide many details about the music including links to other tables like artists, albums, keys, and others. They have the structure shown in Figure 13.

The first two bytes, labeled u_1 , have an unknown purpose; they usually are 24 followed by 00. They are followed by the unexplained two-byte value found in many row types that Mr. Flesniak called $index_shift$, and a four-byte value he called bitmask, although we do not know what the bits mean. The value at bytes 08-0b, $sample_rate$, is the first one we have a solid understanding of: it holds the playback sample rate of the audio file, in samples per second.

	0 1	2 3	4 5	6 7	8 9	a b	c d	e f
00	u_1	i_{shift}	bitn	bitmask		$sample_rate$		ser_id
10	$file_$	$_size$	u	u_2		u_3 u_4		ork_id
20	key	_id	orig_ar	$rtist_id$	labe	l_id	remix	cer_id
30	bitr	rate	track_r	number	ten	npo	genr	e_id
40	albu	m_id	artis	st_id	i	d	$disc_n$	$play_c$
50	year	s_{depth}	dur	u_5	c_{id} r	u_6	u_7	ofs_0
60	ofs_1	ofs_2	ofs_3	ofs_4	ofs_5	ofs_6	ofs_7	ofs_8
70	ofs_9	ofs_{10}	ofs_{11}	ofs_{12}	ofs_{13}	ofs_{14}	ofs_{15}	ofs_{16}
70	ofs_{17}	ofs_{18}	ofs_{19}	ofs_{20}				

Figure 13: Track Row

Bytes $0\mathsf{C}$ - $0\mathsf{f}$ hold the value $composer_id$ which identifies the composer of the track, if known, as a non-zero id value of an Artist row, as discussed in Section 1.3.2. The size of the audio file, in bytes, is found in $file_size$ at bytes 10-13. This is followed by an unknown four-byte value, u_2 , which may be another ID, and two unknown two-byte values, u_3 (about which Mr. Flesniak says "always 19048?") and u_4 ("always 30967?").

If there is cover art for the track, there will be a non-zero value in $artwork_id$ (bytes 1C-1f), identifying the id of an Artwork row, as discussed in section 1.3.3.

If a dominant musical key was identified for the track there will be a non-zero value in key_id (bytes 20-23), which represents the id of a Key row, as discussed in Section 1.3.6. If the track is known to be a remake, the non-zero Artist row id (Section 1.3.2) of the original performer will be found at bytes 24-27 in $original_artist_id$. If there is a known record label for the track, the non-zero value in $label_id$ will link to the id of a Label row id as described in Section 1.3.7. Similarly, if there is a known remixer, there will be a non-zero value in $remixer_id$ (bytes 2C-2f) linking to the id of an Artist row (Section 1.3.2).

The field bitrate at bytes 30-33 stores the playback bit rate of the track, and $track_number$ at bytes 34-37 holds the position of the track within its album. tempo at bytes 38-3b holds the playback tempo of the start of the track in beats per minute, multiplied by 100 (in order to support a precision of $\frac{1}{100}$ BPM). If there is a known genre for the track, there will be a non-zero value in $genre_id$ at bytes 3 C - 3 f, representing the id of a Genre row as discussed in Section 1.3.5.

If the track is part of an album, there will be a non-zero value in $album_id$ at bytes 40-43, and this will be the id of an Album row as discussed in Section 1.3.1. The Artist row id (Section 1.3.2) of the primary performer associated with the track is found in $artist_id$ at bytes 44-47. And the id of the track itself

is found in id at bytes 48-4b. If the album is known to consist of multiple discs, the disc number on which this track is found will be in $disc_number$ at bytes 4C-4d. And the number of times the track has been played is found in $play_count$ (bytes 4e-4f).

The year in which the track was recorded, if known, is in year at bytes 50-51. The sample depth of the track audio file (bits per sample) is in $sample_depth$ at bytes 52-53. The playback time of the track (in seconds, at normal speed) is in duration at bytes 54-55. The purpose of the next two bytes, labeled u_5 , is unknown, they seem to always hold the value 41.

Byte 58, $color_id$ (labeled c_{id} in Figure 13), holds the color assigned to the track in rekordbox, as the id of a Color row (described in Section 1.3.4), or zero if no color has been assigned. Byte 59, rating (labeled r in Figure 13) holds the rating (0 to 5 stars) assigned the track. The next two bytes, labeled u_6 , have an unknown purpose, and seem to always have the value 1. The two bytes after them, labeled u_7 , are also unknown; Mr. Flesniak said "alternating 2 and 3".

The rest of the track row is an array of 21 two-byte offsets that point to DeviceSQL strings. To find the start of the string, add the address of the start of the track row to the offset. The purpose of each string is described in Table 2, and the structure of the strings themselves is explained in Section 1.4.

For convenience, the strings can be accessed as Kaitai Struct instance values with the names shown in the table.

Index	Name	Content
0	$unknown_string_1$	Unknown, so far always empty.
1	texter	Unknown, named by @flesniak.
2	$unknown_string_2$	Unknown, "thought track number, wrong".
3	$unknown_string_3$	Unknown, "Strange strings." ⁴ .
4	$unknown_string_4$	"Strange strings" (as above).
5	messsage	Unknown, named by @flesniak.
6	$kuvo_public$	Empty or "ON". ⁵
7	$autoload_hotcues$	Empty or "ON".6
8	$unknown_string_5$	Unknown.
9	$unknown_string_6$	Unknown, usually empty.
10	$date_added$	When track was added to collection.
11	$release_date$	When track was released.
12	mix_name	Name of the track remix.
13	$unknown_string_7$	Unknown, usually empty.
14	$analyze_path$	File path of track analysis, see Section 2.

Table 2: Track Offset Strings

 $^{^4\}textsc{Often}$ zero length, sometimes low binary values, ASCII 01 or 02 as content.

⁵Apparently used rather than a simple bit flag to control whether the track information is visible on Suyo.

Kuvo.

⁶Apparently used rather than a simple bit flag to control whether hot cues are automatically loaded for the track.

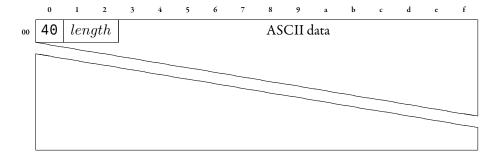


Figure 14: Long ASCII DeviceSQL String

Index	Name	Content
15	$analyze_date$	When track analysis was performed.
16	comment	Track comment assigned by DJ.
17	title	Track title.
18	$unknown_string_8$	Unknown, usually empty.
19	filename	Name of track audio file.
20	$file_path$	File path of track audio.

Table 2: Track Offset Strings

1.4 DeviceSQL Strings

Many row types store string values, sometimes by directly embedding them, but more often by storing an offset to a location elsewhere in the heap. In either case the string itself uses the strange structure described in this section. Strings can be stored in a variety of formats. The first byte of the structure, labeled $length_and_kind$ in the parsed Kaitai Struct, identifies the encoding type and, when the value is odd, also the length (for short ASCII strings), as detailed in Section 1.4.3.

1.4.1 Long ASCII Strings

If $length_and_kind$ has the value 40, it is followed by a two-byte length field, and then followed by that many bytes of ASCII-encoded string data, as shown in Figure 14.

1.4.2 Long UTF-16 Big-Endian Strings

If $length_and_kind$ has the value 90, it is followed by a two-byte length field, and then followed by that many bytes of UTF-16 big-endian encoded string data, as shown in Figure 15.

16

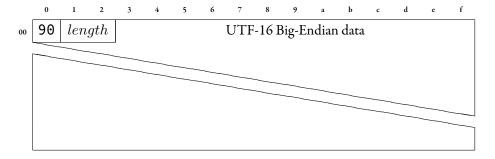


Figure 15: Long UTF-16-BE DeviceSQL String

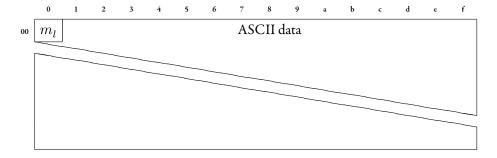


Figure 16: Short ASCII DeviceSQL String

1.4.3 Short ASCII Strings

If $length_and_kind$ has an odd value it is a $mangled_length$, labeled m_l in Figure 16. This means we are dealing with a short ASCII DeviceSQL string. To find the length of the string data (which immediately follows this byte), subtract 1 from $mangled_length$, divide it by 2, and subtract 1 again.

2 Analysis Files

When rekordbox analyzes tracks there is some data that is too big to fit in the database itself. We have already seen some of that (the album art images, and of course the track audio is left in the filesystem as well). The other analysis data is organized into "anlz" files, whose path can be found in the DeviceSQL string pointed to by index 14 in the string offsets found at the end of the corresponding track row (see Table 2 in Section 1.3.10). These files have names like ANLZ0001. DAT and their structure is described in this section.

The files are "tagged type" files, where there is an overall file header section, and then each entry in the file has its own header which identifies the type and length of that section.

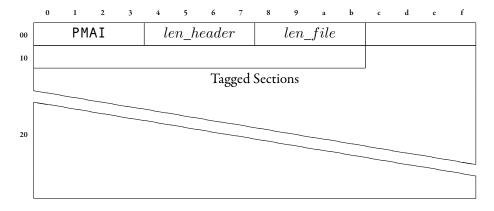


Figure 17: Analysis File Structure

Later player hardware added support for things like colored and more-detailed waveforms. Apparently these were deemed too large to fit in the . DAT files, so another file was introduced, which shares the same base filename as the . DAT file, but uses an extension of . EXT instead. Both kinds of file share the same structure, but different sets of tags can be found in each.

2.1 Analysis File Header

For some reason the analysis files store their numbers in big-endian byte order, the opposite of the export.pdb database file. Field names used in the Figures match the IDs assigned to them in the Kaitai Struct specification⁷, unless that is too long to fit, in which case a subscripted abbreviation is used, and the text will mention the actual struct field name.

The file itself starts with the four-character code PMAI that identifies its format. This file format identifier is followed a four-byte value, len_header (at bytes 04-07) that specifies the length of the file header in bytes. This is followed by another four-byte value, len_file , at bytes 08-0b that specifies the length of the whole file in bytes.

The header seems to usually be 1C bytes long, though we do not yet know the purpose of any of the header values that come after len_file . After the header, the file consists of a series of tagged sections, each with their own four-character code identifying the seciton type, followed by a header and the section content. This overall structure is illustrated in Figure 17, and the structure of the known tag types is described next.

 $^{^7} https://github.com/Deep-Symmetry/crate-digger/blob/master/src/main/kaitai/rekordbox_anlz.ksy$

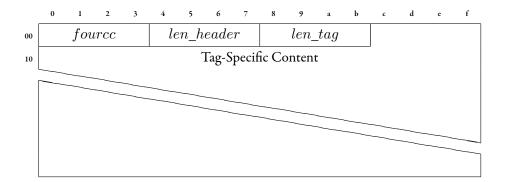


Figure 18: Tagged Section Structure

2.2 Analysis File Sections

The structure of each tagged section has an "envelope" that can be understood even if the internal structure of the section is unknown, making it easy to navigate through the file looking for the section you need. This structure is very similar to the file itself, and is illustrated in Figure 18.

Every section begins with a four-character code, fourcc, identifying its specific structure and content, as described in the sections below. This is followed by a four-byte value, len_header , which specifies how many bytes there are in the section header, and another four-byte value, len_tag , which specifies the length of the entire tagged section (including the header), in bytes. This value can be added to the address of the start of the tag to find the start of the next tag.

There is not much value to len_header . If you study the structure of each type of tagged section, you can see some sense of where the "header-like stuff" ends, and "content-like stuff" begins, and this seems to line up with the value of len_header . But because there are important values in each tag's header, and those always start immediately after len_tag , it is simply easier to ignore the value of len_header , and model the tag body as beginning at byte θC of the tag. To show where the boundary occurs, in the diagrams that follow, values that fall inside the byte range of the header are colored yellow.

2.2.1 Beat Grid Tag

This kind of section holds a list of all beats found within the track, recording their bar position, the time at which they occur, and the tempo at that point. It is identified by the four-character code PQTZ, which may stand for "Pioneer Quantization". It has the structure shown in Figure 19. len_header is 18. The tagspecific content starts with two unknown values, although Mr. Flesniak says that $unknown_2$ seems to always have the value 00800000.

len_beats at bytes 14-17 specifies the number of beats were found in the track, and thus the number of beat entries that will be present in this section. The

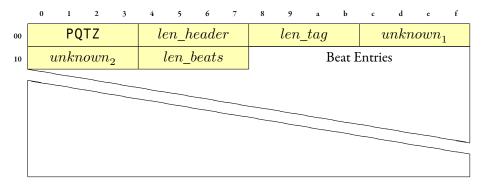


Figure 19: Beat Grid Tag



Figure 20: Beat Grid Beat

beat entries come next, and each has the structure shown in Figure 20.

Each beat entry is eight bytes long. It starts with $beat_number$, a two-byte number (abbreviated b_{num} in Figure 20) which specifies where the beat falls within its measure. So the value is always 1, 2, 3, or 4. This is followed by a two-byte tempo value, which records the track tempo at the point where this beat occurs, in beats per minute multiplied by 100 (to allow a precision of $\frac{1}{100}$ BPM). Finally, there is a four-byte time value, which specifies the time at which this beat would occur, in milliseconds, when playing the track at its normal speed.

As noted above, there will be as many beat entries as len_beats specifies. They continue to the end of the tag.

2.2.2 Cue List Tag

This kind of section holds either a list of ordinary memory points and loops, or a list of hot cues and hot loops. It is identified by the four-character code PCOB, and has the structure shown in Figure 21. len_header is 18.

Note that since the release of the Nexus 2 series of players, there is a newer tag available that contains more information and supports more hot cues, so you should check for that before loading this tag. See Section 2.2.3 for details.

The type value at bytes 0c - 0f determines whether this section holds memory points (if type is 0) or hot cues (if type is 1). The number of cue entries present in the section is reported in len_{cues} at bytes 12 - 13, and we don't yet know the meaning of unk at bytes 10 - 11 or $memory_count$ at bytes 14 - 17. The remainder of the section, from byte 18 through len_tag holds the cue entries them-

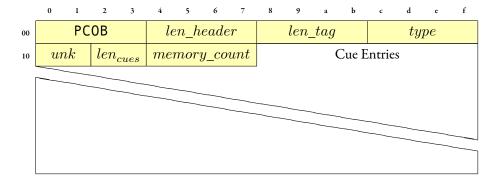


Figure 21: Cue List Tag

	0	1	2	3	4	•	6	7	8	9	a	ь	c	d	e	t
00		PC	PT		$l\epsilon$	en_h	eade	r		len_e	entry	/		$hot_$	_cue	
10		sta	tus		u	nkn	own	1	o_{fi}	irst	o_{la}	ist	t	uni	knou	m_2
20		tir	ne		l	oop_	time	9)			ι	nkn	own	3		
30																

Figure 22: Cue List Entry

selves, with the structure shown in Figure 22.

Each cue entry is 38 bytes long. It is structured as its own miniature tag for unknown reasons, starting with the four-character code PCPT (Pioneer Cue Point?), and its own internal four-byte len_header and len_entry values (1C and 38 respectively).

If the cue is an ordinary memory point, hot_cue at bytes 0C-0f will be zero, otherwise it identifies the number of the hot cue that this entry represents (Hot Cue A is number 1, B is 2, and so on). The status value at bytes 10-13 seems to be a deletion indicator; if it is zero, the entry is ignored. Cues which the players pay attention to have the value 1 here.

The next four bytes have an unknown purpose, but seem to always have the value 00100000. They are followed by two two-byte values, which seem to be for sorting the cues in the proper order in some strange way. $order_first$ at bytes 1a-1b (labeled o_{first} in Figure 22) has the value ffff for the first cue, 0000 for the second, then 2, 3 and on. $order_last$ at bytes 1a-1b (labeled o_{last}) has the value 1 for the first cue, 2 for the second, and so on, but ffff for the last. It would seem that the cues could be perfectly well sorted by just one of these fields, or indeed, by their time values.

The first "non-header" field is type at byte 1C (labeled t in Figure 22), and it

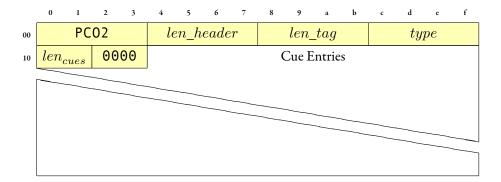


Figure 23: Extended (nxs2) Cue List Tag

specifies whether the entry records a simple position (if it has the value 1) or a loop (if it has the value 2). The next three bytes have an unknown purpose, but seem to always have the values 0003e8, or decimal 1000.

The value *time* at bytes 20-23 records the position of the cue within the track, as a number of milliseconds (representing when the cue would occur if the track is being played at normal speed). If *type* is 2, so this cue stores a loop, then *loop_time* at bytes 24-27 stores the track time in milliseconds at which the player should loop back to *time*.

We do not know what, if anything, is stored in the remaining bytes of the cue entry.

2.2.3 Extended (nxs2) Cue List Tag

This is a variation of the Cue List Tag described in Section 2.2.2 that was introduced with the Nexus 2 players to add support for more than three hot cues with custom color assignments, as well as DJ-assigned comment text for each hot cue and memory point. It also contains the information present in the standard Cue List Tag, so you only need to read one set or the other. Beat Link tries to use the extended tags if they are available, and falls back to using the older ones if they are

Just like the older tag, this kind of section holds either a list of ordinary memory points and loops, or a list of hot cues and hot loops. It is identified by the four-character code PCO2, and has the structure shown in Figure 23. len_header is 14.

The type value at bytes 0C-0f determines whether this section holds memory points (if type is 0) or hot cues (if type is 1). The number of cue entries present in the section is reported in len_{cues} at bytes 10-11, and we don't yet know the meaning of the remaining two header bytes. The remainder of the section, from byte 14 through len_{tag} holds the cue entries themselves, with the structure shown in Figure 24.

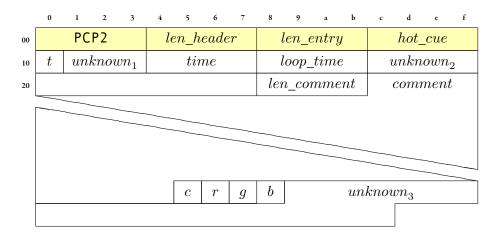


Figure 24: Extended (nxs2) Cue List Entry

Each extended cue entry has a variable length. It is structured as its own miniature tag, starting with the four-character code PCP2, and its own internal four-byte len_header and len_entry values. While len_header has the fixed value 10, len_entry is needed to determine the length of the entry, so the beginning of the next one can be located.

If the cue is an ordinary memory point, *hot_cue* at bytes 0C - 0f will be zero, otherwise it identifies the number of the hot cue that this entry represents (Hot Cue A is number 1, B is 2, and so on).

The *status* flag and mysterious sort order values present in the older cue list entry header are simply absent here.

The first "non-header" field is type at byte 10 (labeled t in Figure 24), and it specifies whether the entry records a simple position (if it has the value 1) or a loop (if it has the value 2). The next three bytes have an unknown purpose, but seem to always have the values 0003e8, or decimal 1000.

The value time at bytes 14-17 records the position of the cue within the track, as a number of milliseconds (representing when the cue would occur if the track is being played at normal speed). If type is 2, so this cue stores a loop, then $loop_time$ at bytes 18-1b stores the track time in milliseconds at which the player should loop back to time.

The next twelve bytes have an unknown purpose, but seem to have the value 00, except for the second byte which seems to have the value 10. They are followed by $len_comment$, which contains the length, in bytes, of the comment field which immediately follows it. If $len_comment$ has a non-zero value, comment will hold the text of the comment, encoded as a UTF-16 Big Endian string with a trailing NUL character. So the length will always be even, and (when non-zero) always at least 4 (a one character comment followed by the trailing NUL).

Immediately after comment (in other words, starting $len_comment + 1c$

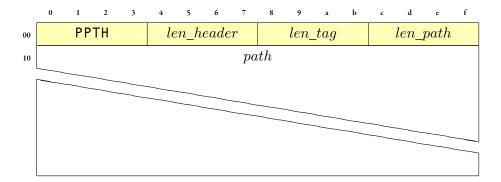


Figure 25: Path Tag

past the start of the entry) there are four one-byte values containing color information. colorCode (labeled c in Figure 24) appears to be a code identifying the color with which rekordbox displays the cue, by looking it up in a table. There have been sixteen codes identified, and their corresponding RGB colors can be found by looking at the findRecordboxColor static method in the Beat Link library's Cuelist class. The next three bytes, $color_red$ (labeled r), $color_green$ (labeled g), and $color_blue$ (labeled b), make up an RGB color specification which is similar, but not identical, to the color that rekordbox displays. We believe these are the values used to illuminate the RGB LEDs in a player that has loaded the cue. When no color is associated with the cue, all four of these bytes have the value 00. We do not know what, if anything, is stored in the remaining bytes of the tag.

2.2.4 Path Tag

This kind of section holds the file path of the audio file for which the track analysis was performed. It is identified by the four-character code PPTH and has the structure shown in Figure 25. len_header is 10.

len_path at bytes 0C - 0f holds the length of the file path value, which makes up the entire tag body. path, which starts at byte 10, is a DeviceSQL string with the structure described in Section 1.4.

2.2.5 **VBR Tag**

This kind of section has not yet been explained, but it is believed to hold an index allowing rapid seeking to particular times within variable-bit-rate tracks. (Without such a structure, it would be necessary to scan the entire file from the beginning to find a frame starting at a particular time, which would be too slow for jumping to memory points or hot cues deep within the track.) What is known of the structure

⁸https://deepsymmetry.org/beatlink/apidocs/

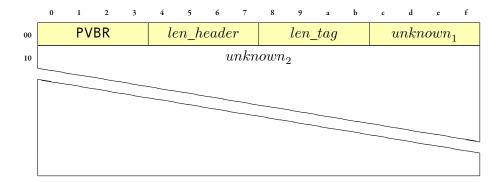


Figure 26: VBR Tag

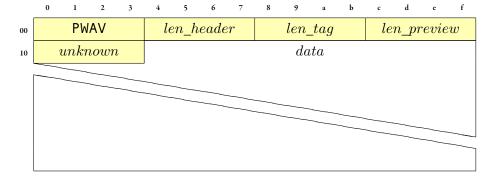


Figure 27: Waveform Preview Tag

is shown in Figure 26. The four-character code that identifies this type of section is PVBR and len_header is 10.

2.2.6 Waveform Preview Tag

This kind of section holds a fixed-width monochrome preview of the track waveform, displayed above the touch strip on original nexus players, providing a birdseye view of the current playback position, and supporting direct needle jump to specific track sections. It is identified by the four-character code PWAV and has the structure shown in Figure 27. len_header is 14.

The purpose of the header bytes 10 - 13 is unknown; they always seem to have the value 00100000. The waveform preview data begins at byte 14 and is 400 (decimal) bytes long. Each byte encodes one vertical pixel-wide column of the waveform preview. The height of the column is represented by the five low-order bits of the byte (so it can range from 0 to 31 pixels high), and the whiteness of the segment is represented by the three high-order bits. Segments with higher values

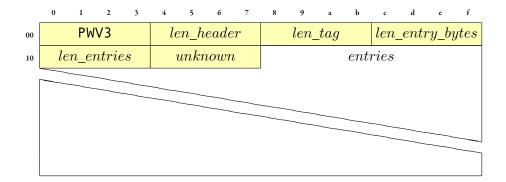


Figure 28: Waveform Detail Tag

in these bits sections are drawn in a less saturated (whiter) shade of blue.

2.2.7 Tiny Waveform Preview Tag

This kind of section holds an even smaller fixed-width monochrome preview of the track waveform, which seems to be displayed on the CDJ-900. It is identified by the four-character code PWV2 but otherwise has the same structure as the larger waveform preview tags shown in Figure 27. len_header is still 14, and header bytes 10-13 also seem to have the value 00100000. The waveform preview data begins at byte 14 and is 100 (decimal) bytes long. Each byte encodes one vertical pixel-wide column of the waveform preview. The height of the column is represented by the four low-order bits of the byte (so it can only range from 0 to 15 pixels high), and no other bits are used.

2.2.8 Waveform Detail Tag

This kind of section holds a variable-width and much larger monochrome preview of the track waveform, which scrolls along while the track plays, giving a detailed glimpse of the neighborhood of the current playback position. Since this is potentially much larger than other analysis elements, and is not supported by older players, it is stored in the extended analyis file (with extension . EXT). It is identified by the four-character code PWV3 and has the structure shown in Figure 28. len_header is 18.

len_entry_bytes identifies how many bytes each waveform detail entry takes up; for this kind of tag it always has the value 1. len_entries specifies how many entries are present in the tag. Each entry represents one half-frame of audio data, and there are 75 frames per second, so for each second of track audio there are 150 waveform detail entries. The purpose of the header bytes 14 - 17 is unknown; they always seem to have the value 00960000. The waveform detail entries begin at byte 18. The interpretation of each byte is the same as for the Waveform Preview data described in Section 2.2.6.

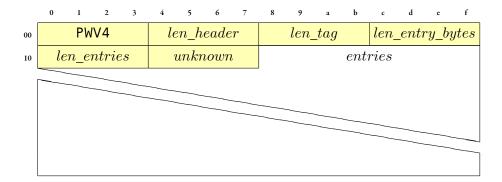


Figure 29: Waveform Color Preview Tag

2.2.9 Waveform Color Preview Tag

This kind of section holds a fixed-width color preview of the track waveform, displayed above the touch strip on nexus 2 players, providing a birds-eye view of the current playback position, and supporting direct needle jump to specific track sections. It is also used in rekordbox itself. This is stored in the extended analyis file (with extension . EXT). It is identified by the four-character code PWV4 and has the structure shown in Figure 29. len_header is 18.

len_entry_bytes identifies how many bytes each waveform preview entry takes up; for this kind of tag it always has the value 6. len_entries specifies how many entries are present in the tag. The purpose of the header bytes 14 - 17 is unknown. The waveform color preview data begins at byte 18 and is 7,200 (decimal) bytes long, representing 1,200 columns of waveform preview information.

The color waveform preview entries are the most complex of any of the waveform tags. See the discussion on Github⁹ for how the analysis was performed. @jan2000 created an audio file containing a 10 second sine wave sweep from 20 Hz to 20 kHz, and analyzed that in rekordbox. The results are represented in Figure 30.

As a summary, the top six stripes plot the values of each six channels of waveform preview information. The first byte of data is the first column of the top stripe, the next byte is the first column of the second stripe, and so on, until we reach the seventh byte, which is the second column of the first stripe.

We are not sure what the top two stripes represent, but they do seem to have an effect on the blue version of the waveform preview, so they somehow encode "whiteness". The next stripe, corresponding to byte 2 of each column, indicates how much sound energy is present in the bottom half of the frequency range (it drops around 10 KHz). The stripe corresponding to byte 3 reflects how much sound energy is present in the bottom third of the frequency range, byte 4 reflects how much sound energy is in the middle of the frequency range, and byte 5 tracks

⁹https://github.com/Deep-Symmetry/dysentery/issues/9

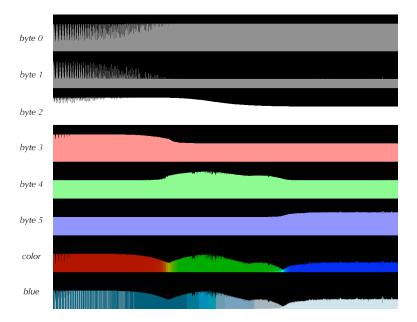


Figure 30: Sine sweep analysis

the sound energy in the top of the frequency range.

The stripe labeled "color" reflect's @jan2000's algorithm for combining bytes 3, 4, and 5 into a color preview, and the bottom stripe is his approach for deriving the blue preview from that and the other two stripes.

The calculations used by Beat Link to build its own color previews can be found in the <code>segmentColor</code> and <code>segmentHeight</code> methods of the <code>Waveform-Preview</code> class¹⁰, and the way they are used to draw the actual graphical representation can be found in the <code>updateWaveform</code> method of the <code>Waveform-PreviewComponent</code> class¹¹. These produce attractive results, but it is certainly possible that refinements can be found in the future.

2.2.10 Waveform Color Detail Tag

This kind of section holds a variable-width and much larger color preview of the track waveform, introduced with the nexus 2 line, which scrolls along while the track plays, giving a detailed glimpse of the neighborhood of the current playback position. This is stored in the extended analyis file (with extension . EXT). It is

¹⁰https://deepsymmetry.org/beatlink/apidocs/org/deepsymmetry/ beatlink/data/WaveformPreview.html

 $^{^{11}} https://deepsymmetry.org/beatlink/apidocs/org/deepsymmetry/beatlink/data/WaveformPreviewComponent.html\\$

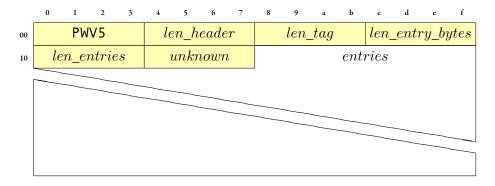


Figure 31: Waveform Color Detail Tag



Figure 32: Waveform color detail segment bits

identified by the four-character code PWV5 and has the structure shown in Figure 31. len_header is 18.

len_entry_bytes identifies how many bytes each waveform detail entry takes up; for this kind of tag it always has the value 2. len_entries specifies how many entries are present in the tag. Each entry represents one half-frame of audio data, and there are 75 frames per second, so for each second of track audio there are 150 waveform detail entries. The purpose of the header bytes 14 - 17 is unknown; they may always have the value 00960305. The color waveform detail entries begin at byte 18.

Color detail entries are much simpler than color preview entries. They consist of three-bit red, green, and blue components and a five-bit height component packed into the sixteen bits of the two entry bytes. Considering each entry as a two-byte big-endian integer, the red component is the three high-order bits. The next three bits are the green component, followed by the three bits of blue intensity, and finally five bits of height. The two low-order bits do not seem to be used. This is shown in Figure 32.

2.2.11 Song Structure Tag

This kind of section is used only in rekordbox Performance Mode, and so does not get exported to external media. If you want to work with song structure data, you will need to find the extended analysis file on the computer hosting rekordbox. Within that file, the section is identified by the four-character code PSSI and has the structure shown in Figure 33. len_header is 20. Many thanks to Michael

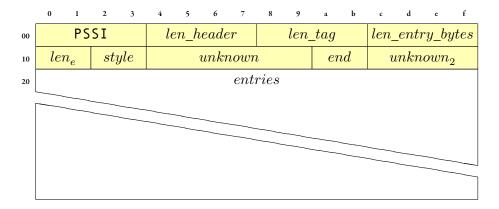


Figure 33: Song Structure Tag

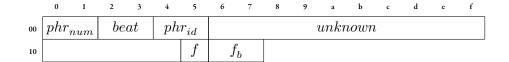


Figure 34: Song Structure Entry

Ganss, https://github.com/mganss, for contributing this analysis.

 len_entry_bytes identifies how many bytes each phrase entry takes up; so far it always has the value 18, so each entry takes twenty four bytes. $len_entries$ at bytes 10– 11 (labeled len_e in Figure 33) specifies how many entries are present in the tag. Each entry represents one recognized phrase. $phrase_style$ at bytes 12– 13 (labeled phrase) specifies the overall type of phrase structure that rekordbox chose to represent the song. The value 1 is an "up-down" style where the main phrases consist of "UP", "DOWN", and "CHORUS", and is represented in rekordbox by white label text. The value 2 is a "bridge-verse" style where the main phrases are labeled "VERSE", "CHORUS", and "BRIDGE", and the label text in rekordbox is black. Style 3 is mostly identical to bridge-verse, except verses 1–3 are labeled "VERSE1" and verses 4–6 are labeled "VERSE2" in rekordbox.

The purpose of the header bytes 14 - 19 is unknown. end_beat at bytes 1a - 1b (labeled end in Figure 33) holds the beat number at which the last recognized phrase ends. The track may continue beyond this, but will mostly be silence from then on.

The final four bytes of the header, 1c - 1f, also have an unknown purpose.

The phrase entries begin at byte 20, and each has the structure shown in Figure 34.

The first two bytes of each song structure entry hold $phrase_number$ (labeled phr_{num} in Figure 34) which numbers each phrase, starting at one and in-

crementing with each entry. That is followed by $beat_number$ (labeled beat), a two-byte value that specifies the beat at which this phrase begins in the track. It continues until either the beat number of the next phrase, or the beat identified by end in the tag header if this is the last entry.

 $phrase_id$ at bytes 4-5 (labeled phr_{id} in Figure 34) specifies what kind of phrase rekordbox has identified here. The interpretation depends on the value of $phrase_style$ in the tag header, as is detailed in Table 3 below.

The purpose of the entry bytes 6-14 is unknown. $fill_in$ at byte 15 (labeled f in Figure 34) is a flag that indicates whether there are fill (non-phrase) beats at the end of the phrase. If it is non-zero, then $fill_in_beat_number$ at bytes 16-17 (labeled f_b) holds the beat number at which the fill begins. When fill-in is present, it is indicated in rekordbox by little dots on the full waveform. The manual says:

[Fill in] is a section that provides improvisational changes at the end of phrase. [Fill in] is detected at the end of Intro, Up, and Chorus (up to 4 beats).

Phrase Style	Phrase ID	Label
1	1	INTRO
1	2	UP
1	3	DOWN
1	5	CHORUS
1	6	OUTRO
2	1	INTRO
2	2	VERSE1
2	3	VERSE2
2	4	VERSE3
2	5	VERSE4
2	6	VERSE5
2	7	VERSE6
2	8	BRIDGE
2	9	CHORUS
2	10	OUTRO
3	1	INTRO
3	2	VERSE1
3	3	VERSE1
3	4	VERSE1
3	5	VERSE2
3	6	VERSE2
3	7	VERSE2
3	8	BRIDGE
3	9	CHORUS

Table 3: Phrase Labels

Phrase Style	Phrase ID	Label
3	10	OUTRO

Table 3: Phrase Labels

Note that because Kaitai Struct does not allow multiple enum values to share the same label, style 3 is modeled in Crate Digger as identical to style 2.

3 Crate Digger

You can find a Java library that can parse the structures described in this research, and that can retrieve them from players' NFS servers, at: https://github.com/deep-symmetry/crate-digger

The project also contains Kaitai Struct specifications for the file structures, which were used to automatically generate Java classes to parse them, and which can be used to generate equivalent code for a variety of other programming languages.

There are also ONC RPC specification files which were similarly used to generate Java classes to communicate with the NFSv2 servers in the players, and which can likely be used to generate structures for other languages as well.

List of Figures

1	File Header	4
2	Table Pointer	4
3	Table Page	6
4	Album Row	8
5	Artist Row with Nearby Name	9
6	Artist Row with Far Name	9
7	Artwork Row	9
8	Color Row	10
9	Genre or Label Row	10
10	Key Row	11
11	Playlist Tree Row	12
12	Playlist Entry Row	12
13	Track Row	13
14	Long ASCII DeviceSQL String	15
15	Long UTF-16-BE DeviceSQL String	16
16	Short ASCII DeviceSQL String	16
17	Analysis File Structure	17
18	Tagged Section Structure	18
19	Beat Grid Tag	19
20	Beat Grid Beat	19

LIST OF TABLES 32

21	Cue List Tag	20
22	Cue List Entry	20
23	Extended (nxs2) Cue List Tag	21
24	Extended (nxs2) Cue List Entry	22
25	Path Tag	23
26	VBR Tag	24
27	Waveform Preview Tag	24
28	Waveform Detail Tag	25
29	Waveform Color Preview Tag	26
30	Sine sweep analysis	27
31	Waveform Color Detail Tag	28
32	Waveform color detail segment bits	28
33	Song Structure Tag	29
34	Song Structure Entry	29
List o	f Tables	
List o		5
	Table Types	-
1	Table Types	-
1 2	Table Types	14
1 2 2	Table Types	14 15

http://deepsymmetry.org