Additions for pre-1921 Latvian orthography

- U+A796 LATIN CAPITAL LETTER G WITH DIAGONAL STROKE
- U+A797 LATIN SMALL LETTER G WITH DIAGONAL STROKE
- U+A798 LATIN CAPITAL LETTER K WITH DIAGONAL STROKE OVERLAY
- U+A799 LATIN SMALL LETTER K WITH DIAGONAL STROKE OVERLAY
  → A743 latin small letter k with diagonal stroke
  "overlay" in the name indicates contrast to U+A743
- U+A79A LATIN CAPITAL LETTER N WITH DIAGONAL STROKE
- U+A79B LATIN SMALL LETTER N WITH DIAGONAL STROKE
- U+A79C LATIN CAPITAL LETTER R WITH DIAGONAL STROKE
- U+A79D LATIN SMALL LETTER R WITH DIAGONAL STROKE
- U+A79E LATIN CAPITAL LETTER S WITH DIAGONAL STROKE
  Also used in pre-1950 Lower Sorbian orthography
- U+A79F LATIN SMALL LETTER S WITH DIAGONAL STROKE
  → A79E latin small letter long s with diagonal stroke
  Also used for Luiseño (California, USA)

Informative notes and cross references to be added for existing characters:

- U+1E9C LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE
  → A79F latin small letter s with diagonal stroke
  Medievalist use for several abbreviations
  Also used in Fraktur types of pre-1921 Latvian orthography
  Also used in Fraktur types of pre-1950 Lower Sorbian orthography
- U+A743 LATIN SMALL LETTER K WITH DIAGONAL STROKE
  → A799 latin small letter k with diagonal stroke overlay
  Medievalist use for abbreviations of "kalendas" and "karta"
  The diagonal stroke crosses the right leg of the k

All these letters are used in the orthography of Latvia used before 1921 [3] [8] [4], see fig. 1 to 8. The LATIN SMALL LETTER S WITH DIAGONAL STROKE also occurs in a long s form (when used within a Fraktur [Blackletter] font); this form is unified with the already encoded U+1E9C LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE.
The encoding of these letters closes a gap regarding Fraktur (Blackletter) types:

According to [3] (which appeared 1880) pp. 227-234, Fraktur was in common use during the 19th century for the following languages: German, Danish, Swedish, Latvian, Finnish, Czech, Slovak. Even as this list may be incomplete, all letters used for the 19th century orthographies of all major European languages are now included in Unicode, with the exception of Latvian.

Encoding the 10 letters proposed here would enable to create Fraktur fonts which cover the complete 19th century and early 20th century usage in Europe.

It should be noted that the proposed letters are not simply presentation forms of some base letters with the overstriking diacritic U+0338 COMBINING LONG SOLIDUS OVERLAY.

All scans show a specific font design for all proposed letters regarding length, angle, and thickness of the stroke, whether the design is straightforward as in fig. 1 or elaborated as in fig. 4, which in no case exactly resemble a solidus (except the Luiseño example, fig. 9).

Moreover, encoding the proposed letters as separate entities is in line with the encoding e.g. of the slashed SENĆOTEN letters [2], which are in fact composites of a base letter and a solidus (as it is common for North American indigenous languages which got their orthographies after the invention of the typewriter). This encoding was accepted without assigning any canonical or compatibility decomposition to the encoded letters, long after the stopping of encoding really precomposed Latin letters. If this encoding model were changed now in a way that slashed letters are regarded decomposable, a third decomposition type had to be added (as the canonical decomposition is stabilized now) for the already encoded slashed letters.

It should also noted that U+0338 is not explicitly devised to generate script elements (like viramas etc. are for South Asian scripts), but is usable as graphical element for overstrikings in critical typescript editions where it really must appear as a solidus which does not change its shape dependent of the letter which it strikes out.

Besides the historic Latvian use, the s with diagonal stroke has more uses:

- It was used in Lower Sorbian orthography until about 1950 (until the current orthography was introduced) [7].
  See fig. 8.
- It is used currently in the orthography of the Luiseño language spoken in California, USA [4].
  See fig. 9.

Regarding the representative glyphs, the stroke is presented in a flat angle, according to the glyphs presented in fig. 1. However, for the small and capital s, this led to a problematic design when using the Roman glyphs (as Latin characters are presented usually in a serifed Roman font in the code tables), while the flat angle stroke is compatible with the Fraktur glyphs which are mostly used by pre-1921 printed Latvian texts. Thus, a more oblique and solidus-like design was selected for the small and capital s, which also is more appropriate for the modern use in Luiseno.

Properties:

\[
\begin{align*}
A796; \text{LATIN CAPITAL LETTER G WITH DIAGONAL STROKE} ; \text{Lu} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A797; \\
A797; \text{LATIN SMALL LETTER G WITH DIAGONAL STROKE} ; \text{Ll} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A796; ; A796 \\
A798; \text{LATIN CAPITAL LETTER K WITH DIAGONAL STROKE OVERLAY} ; \text{Lu} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A799; \\
A799; \text{LATIN SMALL LETTER K WITH DIAGONAL STROKE OEVERLAY} ; \text{Ll} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A798; ; A798 \\
A79A; \text{LATIN CAPITAL LETTER N WITH DIAGONAL STROKE} ; \text{Lu} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79B; \\
A79B; \text{LATIN SMALL LETTER N WITH DIAGONAL STROKE} ; \text{Ll} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79A; ; A79A \\
A79C; \text{LATIN CAPITAL LETTER R WITH DIAGONAL STROKE} ; \text{Lu} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79D; \\
A79D; \text{LATIN SMALL LETTER R WITH DIAGONAL STROKE} ; \text{Ll} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79C; ; A79C \\
A79E; \text{LATIN CAPITAL LETTER S WITH DIAGONAL STROKE} ; \text{Lu} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79F; ; A79E \\
A79F; \text{LATIN SMALL LETTER S WITH DIAGONAL STROKE} ; \text{Ll} & ; 0 ; \text{L} ; ; ; ; ; \text{N} ; ; ; ; ; A79E; ; A79E
\end{align*}
\]
Bibliography:


   Characters. 
   Unicode Document L2/04-170


   http://www.romana-hamburg.de/Fraktur.htm

   In collaboration with Pamela Munro [and others]. Malki Museum Press 1971.

[7] KAULFÜRST, Fabian (Sorabisches Institut / Serbski Institut, Bautzen/Budyšin, Germany): 
   personal communication 2005-09-06.


   (The Not-German [i.e. Latvian] Opitz. Or, Short Guidance to the Latvian Art of Poetry.) 
   Rīga 1697


Acknowledgements:

Special thanks for valuable information to:

Chris Harvey (Indigenous Language Institute, Santa Fe, New Mexico, USA)
Fabian Kaulfürst (Sorabisches Institut / Serbski Institut, Bautzen/Budyšin, Germany)
Prof. Stefan Kessler (Institut für Fremdsprachliche Philologien / Baltistik der Universität Greifswald, Germany)
Gunars Lucans
Andreas Stötzner (Signographisches Institut Andreas Stötzner, Pegau/Sa., Germany)

and all members of the Unicode mailing list who participated in the discussion.
### LETTISCH.

<table>
<thead>
<tr>
<th>Antiqua</th>
<th>Fraktur</th>
<th>Wort</th>
<th>Antiqua</th>
<th>Fraktur</th>
<th>Wort</th>
<th>Antiqua</th>
<th>Fraktur</th>
<th>Wort</th>
</tr>
</thead>
<tbody>
<tr>
<td>A a</td>
<td>A a</td>
<td>a</td>
<td>J j</td>
<td>J j</td>
<td>j</td>
<td>Pj pj</td>
<td>Pj pj</td>
<td>p'</td>
</tr>
<tr>
<td>Å Å</td>
<td>Å Å</td>
<td>a</td>
<td>K k</td>
<td>K k</td>
<td>k</td>
<td>R r</td>
<td>R r</td>
<td>r</td>
</tr>
<tr>
<td>B b</td>
<td>B b</td>
<td>b</td>
<td>K k</td>
<td>K k</td>
<td>k</td>
<td>R r</td>
<td>R r</td>
<td>r</td>
</tr>
<tr>
<td>Bj bj</td>
<td>Bj bj</td>
<td>b</td>
<td>Li li</td>
<td>Li li</td>
<td>l</td>
<td>S t</td>
<td>S t</td>
<td>z</td>
</tr>
<tr>
<td>D d</td>
<td>D d</td>
<td>d</td>
<td>Li li</td>
<td>Li li</td>
<td>l</td>
<td>S t</td>
<td>S t</td>
<td>z</td>
</tr>
<tr>
<td>Da da</td>
<td>Da da</td>
<td>d</td>
<td>M m</td>
<td>M m</td>
<td>m</td>
<td>M m</td>
<td>M m</td>
<td>m</td>
</tr>
<tr>
<td>Doch doch</td>
<td>Bj bj</td>
<td>d</td>
<td>Mj mj</td>
<td>Mj mj</td>
<td>m'</td>
<td>Sch sch</td>
<td>Sch sch</td>
<td>s</td>
</tr>
<tr>
<td>E e</td>
<td>E e</td>
<td>e</td>
<td>N n</td>
<td>N n</td>
<td>n</td>
<td>Ti ti</td>
<td>Ti ti</td>
<td>t</td>
</tr>
<tr>
<td>Ë ë</td>
<td>Ë ë</td>
<td>ë</td>
<td>Ng ng</td>
<td>Ng ng</td>
<td>ng</td>
<td>Tschtch Tschtch</td>
<td>tschtch</td>
<td>tschtch</td>
</tr>
<tr>
<td>Ö Ö</td>
<td>Ö Ö</td>
<td>Ö</td>
<td>Gg g</td>
<td>Gg g</td>
<td>g</td>
<td>N n</td>
<td>N n</td>
<td>n'</td>
</tr>
<tr>
<td>Œ œ</td>
<td>Œ œ</td>
<td>œ</td>
<td>W W</td>
<td>W W</td>
<td>W</td>
<td>U u</td>
<td>U u</td>
<td>u</td>
</tr>
<tr>
<td>H h</td>
<td>H h</td>
<td>h</td>
<td>O o</td>
<td>O o</td>
<td>o</td>
<td>W W</td>
<td>W W</td>
<td>w</td>
</tr>
<tr>
<td>I i</td>
<td>I i</td>
<td>i</td>
<td>O o</td>
<td>O o</td>
<td>o</td>
<td>W W</td>
<td>W W</td>
<td>w</td>
</tr>
<tr>
<td>Ï ñ</td>
<td>Ï ñ</td>
<td>ñ</td>
<td>P p</td>
<td>P p</td>
<td>p</td>
<td>Z z</td>
<td>Z z</td>
<td>z</td>
</tr>
<tr>
<td>Ò ò</td>
<td>Ò ò</td>
<td>ò</td>
<td>R R</td>
<td>R R</td>
<td>r</td>
<td>A a</td>
<td>A a</td>
<td>a</td>
</tr>
</tbody>
</table>

Fig. 1: Scan from [3], p. 231, listing all letters of the Latvian orthography valid at its date of original print (1880), including all letters proposed here (the LATIN SMALL LETTER S WITH DIAGONAL STROKE is found in its round form in the Fraktur column).

### LETTISCH


Sibotenmann. Diejenigen Buchstaben, die keinen Laut bilden, dürfen nicht getrennt werden. Ebenso darf k (das in der alten Orthographie als Drehungszzeich en dient) von dem vorhergehenden Vokal nicht getrennt werden.

Interpunktion. Wie im Deutschen.

Literatur.


76

Fig. 2: Scan from [8], p. 76, showing the small letterforms and documenting the date until the old Latvian orthography was valid (starting in the sixth line of the right column, translating "Since 1921, a new orthography is introduced ").

Special thanks to Andreas Stötzner (Signographisches Institut Andreas Stötzner, Pegau/Sa., Germany), who provided the reference and the scan to the author of this proposal.
### Fig. 3: Scan from [4], p. 186, showing the small letterforms in the table and the capital forms in the text (last line).

- Special thanks to Andreas Stötzner (Signographisches Institut Andreas Stötzner, Pegau/Sa., Germany), who provided the reference and the scan to the author of this proposal.

### Fig. 4: Excerpt from the accompanying PDF file supplied with the TrueType Fraktur (Blackletter) font "MarsFrakturLett" (using a proprietary encoding), downloadable (as of 2008-10-28) from [5] as http://www.romana-hamburg.de/FrakturLett.ZIP

This font shows elaborated forms for all proposed letters.
The Text is in German, discussing rhyming rules for Latvian. The sentence containing the letter series translates: "You also may rhyme the soft [letters] with their hard [counterparts], and the slashed g,k,l,n,r,s,sch with g,k,l,n,r,s,sch, if only this happens not too often and without necessity."

– Special thanks to Prof. Stefan Kessler (Institut für Fremdsprachliche Philologien / Baltistik der Universität Greifswald, Germany) who provided the reference and the scan to the author of this proposal.


– Special thanks to Gunars Lucans who provided the reference to the author of this proposal.
Fig. 7: Scan from [1], p.595, showing a sample text of Latvian (set in Fraktur), showing some of the proposed letters.

Fig. 8: Scan from a Lower Sorbian text of 1905, provided by [3]. As this text is set in Fraktur, it shows the s with diagonal stroke in its long form only.

Fig. 9: Scan from [6], p.121

---

Special thanks to Chris Harvey (Indigenous Language Institute, Santa Fe, NM, USA) who provided the reference and the scan to the author of this proposal.
A. Title: Proposal to encode 10 Latin letters with diagonal stroke

<table>
<thead>
<tr>
<th>2. Requester's name:</th>
<th>Karl Pentzlin</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Requester type: Individual Contribution</td>
<td></td>
</tr>
<tr>
<td>4. Submission date: 2008-10-30</td>
<td></td>
</tr>
<tr>
<td>5. Requester's reference (if applicable):</td>
<td></td>
</tr>
<tr>
<td>6. Choose one of the following: This is a complete proposal: Yes</td>
<td></td>
</tr>
</tbody>
</table>

B. Technical – General

| 1. Choose one of the following: a. This proposal is for a new script (set of characters): No |
|------------------------|------------------|
| Name of the existing block: Latin Extended-D |
| 2. Number of characters in proposal: 10 |
| 3. Proposed category (select one from below - see section 2.2 of P&P document): A-Contemporary |
| B.1-Specialized (small collection) X B.2-Specialized (large collection) |
| C-Major extinct |
| D-Attested extinct |
| E-Minor extinct |
| F-Archaic Hieroglyphic or Ideographic |
| G-Obscure or questionable usage symbols |
| 4. Is a repertoire including character names provided? Yes |
| a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document? Yes |
| b. Are the character shapes attached in a legible form suitable for review? Yes |
| 5. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? Karl Pentzlin |
| If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: http://www.pentzlin.com/proposalfont.zip (more information in the text file included in that archive) |
| 6. References: a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Yes |
| b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes |
| 7. Special encoding issues: Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? No |
| 8. Additional Information: Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard. |
C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?
   - If YES explain: [No]
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
   - If YES, with whom? [see references in the text]
   - If YES, available relevant documents: [see references in the text]
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
   - Reference: [see text]
4. The context of use for the proposed characters (type of use; common or rare)
   - Reference: [see text]
5. Are the proposed characters in current use by the user community?
   - Reference: [see text]
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
   - If YES, is a rationale provided? [Yes]
   - Reference: Keeping in line with other Latin characters
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
   - Preferred
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
   - If YES, is a rationale for its inclusion provided? [No]
   - Reference: [No]
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
   - If YES, is a rationale for its inclusion provided? [No]
   - Reference: [No]
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
    - If YES, is a rationale for its inclusion provided? [No]
    - Reference: [No]
11. Does the proposal include use of combining characters and/or use of composite sequences?
    - If YES, is a rationale for such use provided? [No]
    - Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? [n/a]
12. Does the proposal contain characters with any special properties such as control function or similar semantics?
    - If YES, describe in detail (include attachment if necessary) [No]
13. Does the proposal contain any ideographic compatibility character(s)?
    - If YES, is the equivalent corresponding unified ideographic character(s) identified? [No]
    - If YES, reference: [No]