Additions for pre-1921 Latvian orthography

G  U+A7A0  LATIN CAPITAL LETTER G WITH OBLIQUE STROKE

G  U+A7A1  LATIN SMALL LETTER G WITH OBLIQUE STROKE

K  U+A7A2  LATIN CAPITAL LETTER K WITH OBLIQUE STROKE

k  U+A7A3  LATIN SMALL LETTER K WITH OBLIQUE STROKE

N  U+A7A4  LATIN CAPITAL LETTER N WITH OBLIQUE STROKE

n  U+A7A5  LATIN SMALL LETTER N WITH OBLIQUE STROKE

R  U+A7A6  LATIN CAPITAL LETTER R WITH OBLIQUE STROKE

r  U+A7A7  LATIN SMALL LETTER R WITH OBLIQUE STROKE

S  U+A7A8  LATIN CAPITAL LETTER S WITH OBLIQUE STROKE

· Also used in pre-1950 Lower Sorbian orthography

s  U+A7A9  LATIN SMALL LETTER S WITH OBLIQUE STROKE

→ 1E9C latin small letter long s with diagonal stroke

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

U+1E9C  LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE

→ A7A9 latin small letter s with oblique stroke

· Medievalist use for several abbreviations

· Used in pre-1921 Latvian orthography

· Used in Blackletter types of pre-1950 Lower Sorbian orthography

All these letters are used in the orthography of Latvia used before 1921 [4] [5] [7], see fig. 1 to 6, 9 to 11. The LATIN SMALL LETTER S WITH OBLIQUE STROKE also occurs in a long s form; this form is unified with the already encoded U+1E9C LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE.

Also, the stroked forms of L/l which are part of the pre-1921 Latvian orthography are proposed to be unified with the already encoded U+0141/0142 LATIN CAPITAL/SMALL LETTER WITH STROKE.

The encoding of the letters proposed here closes a gap regarding Blackletter (Fraktur) types:

According to [4] (which appeared 1880) pp. 227-234, Blackletter 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 Blackletter fonts which cover the complete 19th century and early 20th century usage in Europe.

The stroked forms of G/g, K/k, L/l, N/n, R/r (but not the stroked form of S/s/ſ) roughly conform to the modern use of cedilla in the modern Latvian orthography (which is frequently displayed like a comma below for these members, or as a turned comma above in the case of the lowercase g). However, the stroke cannot be simply regarded as a typographic variant of the modern cedilla, as its appearance is
completely different not only in Blackletter (which was the prevalent printing type before 1921 in Latvia),
but also in Roman type (see fig. 1, 9 to 11).

Besides the historic Latvian use, the s with oblique stroke also was used in Lower Sorbian orthography
until about 1950 (until the current orthography was introduced) [6]. – See fig. 7 and 8.

Regarding the representative glyphs, the stroke is presented in a flat angle, according to the glyphs
presented in fig. 1. Therefore, in the names the term "oblique stroke" is used regarding the stroke.

Properties:
A7A0;LATIN CAPITAL LETTER G WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A1;
A7A1;LATIN SMALL LETTER G WITH OBLIQUE STROKE;lL;0;L;;;;;N;;;;A7A0;A7A0
A7A2;LATIN CAPITAL LETTER K WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A3;
A7A3;LATIN SMALL LETTER K WITH OBLIQUE STROKE;lL;0;L;;;;;N;;;;A7A2;A7A2
A7A4;LATIN CAPITAL LETTER N WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A5;
A7A5;LATIN SMALL LETTER N WITH OBLIQUE STROKE;lL;0;L;;;;;N;;;;A7A4;A7A4
A7A6;LATIN CAPITAL LETTER R WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A7;
A7A7;LATIN SMALL LETTER R WITH OBLIQUE STROKE;lL;0;L;;;;;N;;;;A7A6;A7A6
A7A8;LATIN CAPITAL LETTER S WITH OBLIQUE STROKE;Lu;0;L;;;;;N;;;;A7A9;
A7A9;LATIN SMALL LETTER S WITH OBLIQUE STROKE;lL;0;L;;;;;N;;;;A7A8;A7A8

Properties to be changed for already encoded letters: (reference to new uppercase letter)
1E9C;LATIN SMALL LETTER LONG S WITH DIAGONAL STROKE;lL;0;L;;;;;N;;;;A7A8;A7A8

Note:
The properties (including the case mapping) for the triplet s/s/l of the (proposed) U+A7A8/A7A9 and the
already encoded U+1E9C were modeled after the already encoded triplet s/s/l U+1E60/1E61/1E9B
(except that the latter triplet shows a canonical decomposition which the proposed one does not):
1E60;LATIN CAPITAL LETTER S WITH DOT ABOVE;Lu;0;L;0053 0307;;;;N;;;;1E61;
1E61;LATIN SMALL LETTER S WITH DOT ABOVE;lL;0;L;0073 0307;;;;N;;;;1E60;1E60
1E9B;LATIN SMALL LETTER LONG S WITH DOT ABOVE;lL;0;L;017F 0307;;;;N;;;;1E60;1E60

They resemble also the triplet s/s/l (except that this employs a compatibility equivalence for the long s;
however, adding an equivalent compatibility equivalence for the already encoded U+1E9C to the
proposed U+A7A9 was not considered for several reasons, like the use of U+1E9C in medieval context
where it has a distinctive meaning):
0053;LATIN CAPITAL LETTER S;Lu;0;L;;;;;N;;;;0073;
0073;LATIN SMALL LETTER S;lL;0;L;;;;;N;;;;0053;0053
017F;LATIN SMALL LETTER LONG S;lL;0;L;<compat> 0073;;;;N;;;;0053;0053
Bibliography:

[6] KAULFÜRST, Fabian (Sorbisches Institut / Serbski Institut, Bautzen/Budyšin, Germany):
(The Not-German [i.e. Latvian] Opitz. Or, Short Guidance to the Latvian Art of Poetry.)
Rīga 1697
[9] Agenda. sa tu evangelisku zerkwu we Pschuskej. Worjejce (Hoyerswerda; Germany), 1898.

Acknowledgements:

The proposal was mainly compiled by Karl Pentzlin.

Special thanks for valuable information to:

Fabian Kaulfürst (Sorbisches Institut / Serbski Institut [Sorbian Institute], Bautzen/Budyšin, Germany)
Prof. Stefan Kessler (Institut für Fremdsprachliche Philologien / Baltistik der Universität Greifswald, Germany)
Gunars Lucans
Sonja Pentzlin
Andreas Stötzner (Signographisches Institut Andreas Stötzner, Pegau/Sa., Germany) for some of the samples, for a thorough discussion, and for providing the font used for the representative glyphs
and all members of the Unicode mailing list who participated in the discussion of earlier drafts.
Fig. 1: Scan from [4], 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 OBLIQUE STROKE is found in its round form in the "Fraktur" (Blackletter) column).

Fig. 2: Scan from [7], 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...") – Scan and reference provided by Andreas Stötzer, Signographisches Institut.
Mehrere unserer Laute kommen im Lettischen nicht vor, wie ch ö ü, andere werden umschrieben, z. B. q = ëv, x = ëj, y = ë, ã = e. Auch h wird als selbständiger Buchstabe nicht gebraucht, sondern nur beim ëh und als Dehnzeichen bei Selbstlauten, c nur in der neuen Rechtschreibung und als Hilfsbuchstabe bei ëh.


Fig. 3: Scan from [5], p. 186, showing the small letterforms in the table and the capital forms in the text (last line). – Scan and reference provided by Andreas Stötzer, Signographisches Institut
Fig. 4: Scan from [8], p.19, showing all lower case forms of the 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."

Scan and reference provided by Prof. Stefan Kessler, Universität Greifswald.

Fig. 5: Scan from [10], p.12, showing some of the proposed letters.

Note: This example shows ligatures of two long s with stroke (line starting with "2.") and of two l with stroke (second last line), employing a common stroke between the two letters. Such ligatures are common in 18th century prints but we have found no such ligatures in 19th century prints. We regard such ligatures as presentation forms, like other common Blackletter ligatures like "ch".

Reference provided by Gunars Lucans.
Latvijai tāpat kā viņajai kreisijai viņa vērsība bija jāgrešs už karo-tušu, tādēļ marbuht viņš jebkāds lauzināts polīta neatņemts. Tomēr jāizstāda, ka viņš ir jau jaukākās, jūtas būtīties teksta.

Fig. 6: Scan from [1], p.595, showing a sample text of Latvian (set in Fraktur), showing some of the proposed letters.

Fig. 7: Scan from a Lower Sorbian text of 1905. As this text is set in Blackletter, it shows the s with oblique stroke in its long form only. – Scan provided by Fabian Kaulfürst

Fig. 8: Scan from [9], providing a Lower Sorbian example from 1898, showing also the upper case of the S with oblique stroke. – Scan and reference provided by Fabian Kaulfürst
Fig. 9: Scan from [2] (p.48), providing a Latvian example showing some of the proposed letters, printed in Antiqua type.

Fig. 10: Another scan from [2] (p.48), providing a Latvian example showing some of the proposed letters, printed in Antiqua type.

Fig. 11: Scan from [3] (a Latvian grammar in German), showing Latvian examples in Antiqua type.
### Proposal Summary Form to Accompany Submissions

For Additions to the Repertoire of ISO/IEC 10646


Please ensure you are using the latest Form from [http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html](http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html) for latest Roadmaps.

**A. Administrative**

1. **Title:** Proposal to encode 10 Latin letters for pre-1921 Latvian orthography

2. **Requester's name:** German NB

3. **Requester type (Member body/Liaison/Individual contribution):** Member Body Contribution 2009-03-26, revised 2009-04-30

4. **Submission date:** 2009-03-26, revised 2009-04-30

5. **Requester's reference (if applicable):**

6. **Choose one of the following:**
   - This is a complete proposal: Yes
   - (or) More information will be provided later:

---

**B. Technical – General**

1. **Choose one of the following:**
   - a. This proposal is for a new script (set of characters): No
   - b. The proposal is for addition of character(s) to an existing block: Yes

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)
   - 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

5. **Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?** Andreas Stötzer (Signographisches Institut, Pegau/Sa., Germany)

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](http://www.unicode.org) 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  *(This proposal is a rework of L2/08-405 which was not decided at UTC #117 of Nov. 2008)*  
   - 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.)?  
   - Yes

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?  
   - Yes

4. The context of use for the proposed characters (type of use; common or rare)  
   - see text

5. Are the proposed characters in current use by the user community?  
   - see text

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?  
   - Yes

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?  
   - No

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?  
   - No

10. Can any of the proposed characters be considered to be similar (in appearance or function) to an existing character?  
    - No

11. Does the proposal include use of combining characters and/or use of composite sequences?  
    - No

12. Does the proposal contain characters with any special properties such as control function or similar semantics?  
    - No

13. Does the proposal contain any ideographic compatibility character(s)?  
    - No