

The cmsrb package

Uroš Stefanović*

March 30, 2020 v4.0

1 Why cmsrb?

The `cm-super` package provides great support for Cyrillic script in various languages, but there's a problem with italic variants of some letters for Serbian and Macedonian. The `cmsrb` package includes the correct shapes for italic letters `\cyrg`, `\cyrd`, `\cyrp`, `\cyrt` and letter `\cyrb`. This package also has some improvements in letters and accents used in Serbian language.

2 Package Features

The `cmsrb` fonts are the extensions of the Computer Modern fonts for Serbian and Macedonian languages for \TeX (\LaTeX). Support files are provided for `T1`, `TS1`, `T2A`, `X2` and `OT2` encodings, but also for the experimental `ECMSRB1` and `ECMSRB2` encodings. Supported font styles are serif, sans-serif and mono, for size 10pt.

This package is very simple to use: just put

```
\usepackage{cmsrb}
```

in preamble of the document.

| cmsrb | cm-super | cmsrb | cm-super | cmsrb | cm-super |
|--------------|--------------|--------------|--------------|-----------|-----------|
| <i>đīgūū</i> | <i>đzđnm</i> | <i>đīgūū</i> | <i>đzđnm</i> | đđ | đđ |
| đđđđ | đđđđ | āāŃ | āāŃ | <i>đđ</i> | <i>đđ</i> |

Table 1: Some `cm-super` and `cmsrb` differences.

3 OT2 Features

The `OT2` encoding is modified to match the rules of Latin to Cyrillic transcription in Serbian language. Therefore, the `ts`, `kh`, `ch` and similar ligatures are removed from the encoding (see Table 2).

In `OT2` encoding Macedonian letters `Ǧ` and `Ǧ̄` are added, with Montenegrin letters `Ć` and `Ć̄`, and also some orthographic symbols are included.

However, it is not recommended to use the `OT2` encoding; better choice for Serbian language is `T2A` encoding and `utf8` input. One can use `OT2` encoding if they already have a document written in Latin—in which case it would be easy to transcribe it to Cyrillic.

Example 1:

```
\documentclass{article}
\usepackage{cmsrb}
\usepackage[OT2,T1]{fontenc}
\usepackage[serbian]{babel}
```

*urostajms@gmail.com

| Input | Output | Input | Output | Input | Output | Input | Output |
|----------|--------|----------|--------|----------|--------|----------|--------|
| A | А | B | Б | C | Ц | D | Д |
| E | Е | F | Ф | G | Г | H | Х |
| I | И | J | Ј | K | К | L | Л |
| M | М | N | Н | O | О | P | П |
| Q | Ч | R | Р | S | С | T | Т |
| U | У | V | В | W | Ш | X | Ш |
| Y | Ы | Z | З | \# | Ђ | | |
| a | а | b | б | c | ц | d | д |
| e | е | f | ф | g | г | h | х |
| i | и | j | ј | k | к | l | л |
| m | м | n | н | o | о | p | п |
| q | ч | r | р | s | с | t | т |
| u | у | v | в | w | ш | x | ш |
| y | ы | z | з | + | ђ | | |
| C1 | Ћ | D1 | Ђ | D2 | Џ | D3 | Ѕ |
| E0 | Ě | E1 | Ǽ | E2 | Є | I0 | Й |
| I1 | І | J1 | Ј | J2 | Ю | L1 | Љ |
| N0 | № | N1 | Њ | P1 | Ђ | P2 | Ђ |
| Z1 | Ж | \v {C} | Ч | \'C | Ћ | \DJ | Ђ |
| \v {S} | Ш | \v {Z} | Ж | LJ | Љ | Lj | Љ |
| NJ | Њ | Nj | Њ | D\v {Z} | Џ | D\v {z} | Џ |
| c1 | ћ | d1 | ђ | d2 | џ | d3 | ѕ |
| e0 | ě | e1 | ǽ | e2 | ё | i0 | й |
| i1 | і | j1 | ј | j2 | ю | l1 | љ |
| \i | ı | n1 | њ | p1 | ь | p2 | џ |
| z1 | ж | \v {c} | ч | \'c | ћ | \dj | ђ |
| \v {s} | ш | \v {z} | ж | lj | љ | | |
| nj | њ | | | d\v {z} | џ | | |
| \char 20 | Ѕ | \char 21 | Ѕ | \char 28 | v | \char 29 | ё |
| \'G | Ѓ | \'K | Ѓ | \'g | ѓ | \'k | ќ |
| \'S | Ѓ | \'Z | Ѓ | \'s | с | \'z | з |

Table 2: OT2 encoding in cmsrb package.

```

\newcommand{\test}%
{Ljubazni fenjerd\v zija \v ca\dj avog lica ho\'ce da mi poka\v ze \v stos.}
\begin{document}
\test \\
\fontencoding{OT2}\selectfont \test \\
Акценти: \'a\'a\C a\f a\=a\^a"a\u a
\end{document}

```

Result 1:

Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.
Љубазни фењерџија чађавог лица хоће да ми покаже штос.
Акценти: áàâãäå

Example 2:

```

\documentclass{article}
\usepackage{cmsrb}
\usepackage[OT2,T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[serbian]{babel}
\newcommand{\test}%
{Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.}
\begin{document}
\textit{\test} \\
\fontencoding{OT2}\selectfont \textit{\test}
\end{document}

```

Result 2:

Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.
Љубазни фењерџија чађавог лица хоће да ми покаже штос.

4 T1 Features

The T1 encoding now support conversion from Cyrillic to Latin script (see Example 3).

On the position 208 now is the letter Đ (U+0110) instead of the letter Đ (U+00D0), which enable copying and searching in the Serbian or Croatian Latin text.

Example 3:

```

\documentclass{article}
\usepackage{cmsrb}
\usepackage[T2A,T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[serbian]{babel}
\begin{document}
Љубазни фењерџија чађавог лица хоће да ми покаже штос.
\end{document}

```

Result 3:

Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.

5 T2A Features

The Cyrillic letters J and j are not supported in the T2A encoding, so they are placed instead of the Latin letters J and j. This give us the better results in the Cyrillic text, but that also means this Latin letter are not present in the T2A encoding, so be carefull during searching and copying Latin text in this encoding.

The letters F, K, Ž, Č, f, k, ž, č are removed from the T2A encoding, and on their places now are the letters Ā, Ķ, Ž, Č, f, k, ž, č.

Example 4:

```
\documentclass{article}
\usepackage{cmsrb}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[serbianc]{babel}
\DeclareTextSymbolDefault{\dj}{T1}
\begin{document}
\textit{Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.} \\
Љубазни фењерџија чађавог лица хоће да ми покаже штос.} \\
Акценти: \'{u}\'c\{u}\f{u}\={u}\^{u}\\"{u}\u{u}
\end{document}
```

Result 4:

Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.
Љубазни фењерџија чађавог лица хоће да ми покаже штос.
Акценти: иййййййй

Example 5:

```
\documentclass{article}
\usepackage{cmsrb}
\usepackage[T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[serbianc]{babel}
\begin{document}
\textit{абвгдђђежз\'}{з}сијкљмњњопрс\'}{с}тћкуфхццш}
\end{document}
```

Result 5:

абвгдђђежззсујкљмњњопрсџћкуфхццш

6 TS1 Features

In the TS1 encoding we can find the letter Đ (U+00D0), which is removed from the T1 encoding. Also, in the TS1 encoding now are the letters б, г, д, п, т with Russian italic shape.

Example 6:

```
\documentclass{article}
\usepackage{cmsrb}
\usepackage[TS1,T2A]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[serbianc]{babel}
\begin{document}
\textit{бгдпт} \fontencoding{TS1}\selectfont \textit{бгдпт \DH}
\end{document}
```

Result 6:*бигүү бэднм Д*

7 Math

The `cmsrb` package load default Computer Modern math fonts, but load only fonts with size lower or equal 10pt, for better visual effect. If the option `math` is used, then the package `cmupint` will be loaded, and the integral sign will be changed from \int to \int (upright integral is traditionally used in Serbian language, with `\limits` option). Also, in Serbian language is better to use the symbols \leq and \geq than the symbols \leq and \geq . If the `math` option is used the package `amssymb` will be loaded and this symbols will be changed (also $\not\leq$ and $\not\geq$ will be used).

We use this option like this:

```
\usepackage[math]{cmsrb}
```

Example 7:

```
\documentclass{article}
\usepackage{cmsrb}
\newcommand{\ud}{\, \mathrm{d}}
\begin{document}
$$ \int_0^1 e^x \ud x \geq 0 $$
\end{document}
```

Result 7:

$$\int_0^1 e^x dx \geq 0$$

Example 8:

```
\documentclass{article}
\usepackage[math]{cmsrb}
\newcommand{\ud}{\, \mathrm{d}}
\begin{document}
$$ \int_0^1 e^x \ud x \geq 0 $$
\end{document}
```

Result 8:

$$\int_0^1 e^x dx \geq 0$$

8 Experimental encodings

The encodings `ECMSRB1` and `ECMSRB2` are created for this package only. They are similar to `OT2` encoding, but they contain specific ligatures and accents (see Table 3). For the complete list of characters used in this encodings see Tables on Pages 8 and 9. The purposes of this encodings are writing Serbian text using only ASCII characters, and making simple conversions from Latin to Cyrillic.

Example 9:

| Input | ECMSRB1 | ECMSRB2 | Input | ECMSRB1 | ECMSRB2 | Ulaz | ECMSRB1 | ECMSRB2 |
|-------|---------|---------|-------|---------|---------|------|---------|---------|
| CX | Č | Ч | CY | Ć | Ћ | DX | DŽ | Џ |
| DY | Đ | Ђ | EY | Ě | Ђ | GY | GY | Ѓ |
| KY | KY | Ќ | LJ | LJ | Љ | LY | LJ | Љ |
| NJ | NJ | Њ | NY | NJ | Њ | SX | Š | Ш |
| SXY | Š | Ї | ZX | Ž | Ж | ZY | Ž | S |
| ZXY | Ž | Ž | \EZH | Ž | S | | | |
| Cx | Č | Ч | Cy | Ć | Ћ | Dx | Dž | Џ |
| Dy | Đ | Ђ | Ey | Ě | Ђ | Gy | Gy | Ѓ |
| Ky | Ky | Ќ | Lj | Lj | Љ | Ly | Lj | Љ |
| Nj | Nj | Њ | Ny | Nj | Њ | Sx | Š | Ш |
| Sxy | Š | Ї | Zx | Ž | Ж | Zy | Ž | S |
| Zxy | Ž | Ž | | | | | | |
| cx | č | ч | cy | ć | ћ | dx | dž | џ |
| dy | đ | ђ | ey | ě | ђ | gy | gy | ѓ |
| ky | ky | ќ | lj | lj | љ | ly | lj | љ |
| nj | nj | њ | ny | nj | њ | sx | š | ш |
| sxy | š | џ | zx | ž | ж | zy | ž | s |
| zxy | ž | ž | \ezh | ž | s | | | |
| \'a | à | à | \'a | á | á | \^a | â | â |
| \T a | ã | ã | \D a | ä | ä | \H a | ǎ | ǎ |
| \r a | â | â | \R a | ȁ | ȁ | \v a | ǻ | ǻ |
| \u a | ǻ | ǻ | \U a | ǻ | ǻ | \=a | ā | ā |
| \b a | ȁ | ȁ | \.a | ȁ | ȁ | \d a | ȁ | ȁ |
| \c a | ȁ | ȁ | \k a | ȁ | ȁ | \"a | ǻ | ǻ |
| \C a | ǻ | ǻ | \~a | â | â | \f a | â | â |

Table 3: Encodings ECMSRB1 and ECMSRB2 for the package cmsrb.

```

\documentclass{article}
\usepackage{cmsrb}
\usepackage[ECMSRB1]{fontenc}
\usepackage[serbian]{babel}
\begin{document}
Ljubazni fenjerdxija cxadyavog lica hocye da mi pokazxe sxtos.
\end{document}

```

Result 9:

Ljubazni fenjerdžija čađavog lica hoće da mi pokaže štos.

Example 10:

```

\documentclass{article}
\usepackage{cmsrb}
\usepackage[ECMSRB2]{fontenc}
\usepackage[serbian]{babel}
\begin{document}
Ljubazni fenjerdxija cxadyavog lica hocye da mi pokazxe sxtos.
\end{document}

```

Result 10:

Љубазни фењерџија чађавог лица хоће да ми покаже штос.

9 Version history

4.0

- All options are removed, the option `math` is added.
- Fixed encodings `T1`, `T2A`, `TS1`, `OT2`.
- Letter δ is significantly improved in all styles.
- Experimental encodings `ECMSRB1` and `ECMSRB2` are added.

3.1

- Bugs fixed.
- Options `noint` and `nosymb` are active.
- Symbols \leq , \geq , $\not\leq$, $\not\geq$ are changed.

3.0

- Letter `b` is modified.
- Cyrillic to Latin conversion is added.
- Math fonts are changed (`nomath` option is active).
- Upright integral sign is added.

2.0

- Added support for the letters \acute{r} , \acute{k} , \acute{c} , \acute{s} .

1.1

- The separate `.map` files merged into one.

A Encoding ECMSRB1

| | '0 | '1 | '2 | '3 | '4 | '5 | '6 | '7 | |
|------|-------|-------|-------|-------|-------|--------|--------|--------|-----|
| '00x | ` 0 | ^ 1 | ^ 2 | ~ 3 | ¨ 4 | ˜ 5 | ° 6 | ˘ 7 | ~0x |
| '01x | ˘ 8 | - 9 | · 10 | ˘ 11 | ˘ 12 | ∅ 13 | < 14 | > 15 | |
| '02x | “ 16 | ” 17 | „ 18 | « 19 | » 20 | — 21 | — 22 | € 23 | ~1x |
| '03x | 0 24 | 1 25 | J 26 | ˘ 27 | ^ 28 | § 29 | fi 30 | fl 31 | |
| '04x | ˘ 32 | ! 33 | " 34 | # 35 | \$ 36 | % 37 | & 38 | ' 39 | ~2x |
| '05x | (40 |) 41 | * 42 | + 43 | , 44 | - 45 | . 46 | / 47 | |
| '06x | 0 48 | 1 49 | 2 50 | 3 51 | 4 52 | 5 53 | 6 54 | 7 55 | ~3x |
| '07x | 8 56 | 9 57 | : 58 | ; 59 | < 60 | = 61 | > 62 | ? 63 | |
| '10x | @ 64 | A 65 | B 66 | C 67 | D 68 | E 69 | F 70 | G 71 | ~4x |
| '11x | H 72 | I 73 | J 74 | K 75 | L 76 | M 77 | N 78 | O 79 | |
| '12x | P 80 | Q 81 | R 82 | S 83 | T 84 | U 85 | V 86 | W 87 | ~5x |
| '13x | X 88 | Y 89 | Z 90 | [91 | \ 92 |] 93 | ^ 94 | _ 95 | |
| '14x | ' 96 | a 97 | b 98 | c 99 | d 100 | e 101 | f 102 | g 103 | ~6x |
| '15x | h 104 | i 105 | j 106 | k 107 | l 108 | m 109 | n 110 | o 111 | |
| '16x | p 112 | q 113 | r 114 | s 115 | t 116 | u 117 | v 118 | w 119 | ~7x |
| '17x | x 120 | y 121 | z 122 | { 123 | 124 | } 125 | ~ 126 | - 127 | |
| '20x | Č 128 | Ć 129 | Đ 130 | Š 131 | Ž 132 | Š 133 | Ž 134 | Ž 135 | ~8x |
| '21x | 136 | 137 | Ě 138 | 139 | 140 | Lj 141 | Nj 142 | Dž 143 | |
| '22x | č 144 | ć 145 | đ 146 | š 147 | ž 148 | š 149 | ž 150 | ž 151 | ~9x |
| '23x | 152 | 153 | ě 154 | 155 | 156 | lj 157 | nj 158 | dž 159 | |
| '24x | 160 | 161 | 162 | 163 | 164 | 165 | 166 | 167 | ~Ax |
| '25x | 168 | 169 | № 170 | × 171 | ÷ 172 | LJ 173 | NJ 174 | DŽ 175 | |
| | ~8 | ~9 | ~A | ~B | ~C | ~D | ~E | ~F | |

B Encoding ECMSRB2

| | '0 | '1 | '2 | '3 | '4 | '5 | '6 | '7 | |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-----|
| '00x | ` 0 | ´ 1 | ^ 2 | ~ 3 | ¨ 4 | ˘ 5 | ° 6 | ˘ 7 | ~0x |
| '01x | ˘ 8 | ¯ 9 | · 10 | ¸ 11 | € 12 | ∞ 13 | < 14 | > 15 | |
| '02x | “ 16 | ” 17 | „ 18 | « 19 | » 20 | – 21 | — 22 | € 23 | ~1x |
| '03x | 0 24 | 1 25 | 2 26 | ~ 27 | ^ 28 | § 29 | 30 | 31 | |
| '04x | ˘ 32 | ! 33 | " 34 | # 35 | \$ 36 | % 37 | & 38 | ' 39 | ~2x |
| '05x | (40 |) 41 | * 42 | + 43 | , 44 | - 45 | . 46 | / 47 | |
| '06x | 0 48 | 1 49 | 2 50 | 3 51 | 4 52 | 5 53 | 6 54 | 7 55 | ~3x |
| '07x | 8 56 | 9 57 | : 58 | ; 59 | < 60 | = 61 | > 62 | ? 63 | |
| '10x | @ 64 | А 65 | Б 66 | Ц 67 | Д 68 | Е 69 | Ф 70 | Г 71 | ~4x |
| '11x | X 72 | И 73 | Ј 74 | К 75 | Л 76 | М 77 | Н 78 | О 79 | |
| '12x | П 80 | Ю 81 | Р 82 | С 83 | Т 84 | У 85 | В 86 | Я 87 | ~5x |
| '13x | Ъ 88 | Ь 89 | З 90 | [91 | \ 92 |] 93 | ^ 94 | _ 95 | |
| '14x | ‘ 96 | а 97 | б 98 | ц 99 | д 100 | е 101 | ф 102 | г 103 | ~6x |
| '15x | х 104 | и 105 | ј 106 | к 107 | л 108 | м 109 | н 110 | о 111 | |
| '16x | п 112 | ю 113 | р 114 | с 115 | т 116 | у 117 | в 118 | я 119 | ~7x |
| '17x | ъ 120 | ь 121 | з 122 | { 123 | 124 | } 125 | ~ 126 | - 127 | |
| '20x | Ч 128 | ћ 129 | Ѣ 130 | Ш 131 | Ж 132 | Ѓ 133 | Ѕ 134 | Ѕ 135 | ~8x |
| '21x | Ѓ 136 | Ќ 137 | Ђ 138 | 139 | 140 | Љ 141 | Њ 142 | Џ 143 | |
| '22x | ч 144 | ћ 145 | ђ 146 | ш 147 | ж 148 | ѓ 149 | ѕ 150 | ѕ 151 | ~9x |
| '23x | ѓ 152 | ќ 153 | ђ 154 | 155 | 156 | љ 157 | њ 158 | џ 159 | |
| '24x | 160 | 161 | 162 | 163 | 164 | 165 | 166 | 167 | ~Ax |
| '25x | 168 | 169 | № 170 | × 171 | ÷ 172 | 173 | 174 | 175 | |
| | ~8 | ~9 | ~A | ~B | ~C | ~D | ~E | ~F | |