

Greek characters supported by *greek-fontenc*

Günter Milde

August 16, 2023

Abstract

This document lists all Greek characters that are supported by *greek-fontenc*. It also tests the outcome of \MakeUppercase and \MakeLowercase for different input variants (literal input, LICR¹ macro, Latin transliteration).

The document is compiled with pdfLaTeX, format version 2020-10-01 patch-level 4, and the L3 programming layer from 2021-01-09.

The *babel-greek* version is “2023/06/23 1.13.3 Greek support for the babel system”.

The \greekfontencoding is LGR and the current language variant is *polutoniko*.

Contents

1 Input Variants

The columns represent the following input variants:

1. literal character
2. named accent macro + LICR
3. short accent macro + LICR
4. short accent macro + transliteration
5. accent character + transliteration

For composite diacritics, there are additional columns for short macros in reversed order and with the second diacritic as accent macro.

2 Greek and Coptic

Characters from the Greek and Coptic Unicode block that are supported by *greek-fontenc* and *greek-inputenc*:

x0374	''	''	''
x0375	''	''	''
x037A	⠄⠄⠄⠄	⠄⠄⠄⠄	⠄⠄⠄⠄
0x37E	; ; ; ; ;	; ; ; ; ;	; ; ; ; ;
x0384	⠄⠄⠄⠄	⠄⠄⠄⠄	⠄⠄⠄⠄

¹LaTeX internal character representation

x0385
x0386	'A 'A 'A 'A 'A	A A A A A	á á á á á
x0387
x0388	'E 'E 'E 'E 'E	E E E E E	é é é é é
x0389	'H H H H 'H	H H H H H	ń ń ń ń ń
x038A	I I I I I	I I I I I	í í í í í
x038C	'O 'O 'O 'O 'O	O O O O O	ó ó ó ó ó
x038E	'T 'T 'T 'T 'T	T T T T T	ú ú ú ú ú
x038F	'Ω 'Ω 'Ω 'Ω 'Ω	Ω Ω Ω Ω Ω	á á á á á
x0390	í í í í í í í í í í	I I I I I I I I I I	í í í í í í í í í í
x0391	A A A A A	A A A A A	α α α α α
x0392	B B B B B	B B B B B	β β β β β
x0393	Γ Γ Γ Γ Γ	Γ Γ Γ Γ Γ	γ γ γ γ γ
x0394	Δ Δ Δ Δ Δ	Δ Δ Δ Δ Δ	δ δ δ δ δ
x0395	E E E E E	E E E E E	ε ε ε ε ε
x0396	Z Z Z Z Z	Z Z Z Z Z	ζ ζ ζ ζ ζ
x0397	H H H H H	H H H H H	η η η η η
x0398	Θ Θ Θ Θ Θ	Θ Θ Θ Θ Θ	θ θ θ θ θ
x0399	I I I I I	I I I I I	í í í í í
x039A	K K K K K	K K K K K	κ κ κ κ κ
x039B	Λ Λ Λ Λ Λ	Λ Λ Λ Λ Λ	λ λ λ λ λ
x039C	M M M M M	M M M M M	μ μ μ μ μ
x039D	N N N N N	N N N N N	ν ν ν ν ν
x039E	Ξ Ξ Ξ Ξ Ξ	Ξ Ξ Ξ Ξ Ξ	ξ ξ ξ ξ ξ
x039F	O O O O O	O O O O O	ο ο ο ο ο
0x3A0	Π Π Π Π Π	Π Π Π Π Π	π π π π π
0x3A1	P P P P P	P P P P P	ρ ρ ρ ρ ρ
0x3A3	Σ Σ Σ Σ Σ	Σ Σ Σ Σ Σ	ς ς ς ς ς
0x3A4	T T T T T	T T T T T	τ τ τ τ τ
0x3A5	Υ Υ Υ Υ Υ	Υ Υ Υ Υ Υ	υ υ υ υ υ
0x3A6	Φ Φ Φ Φ Φ	Φ Φ Φ Φ Φ	φ φ φ φ φ
0x3A7	X X X X X	X X X X X	χ χ χ χ χ
0x3A8	Ψ Ψ Ψ Ψ Ψ	Ψ Ψ Ψ Ψ Ψ	ψ ψ ψ ψ ψ
0x3A9	Ω Ω Ω Ω Ω	Ω Ω Ω Ω Ω	ω ω ω ω ω
0x3AA	í í í í í	í í í í í	í í í í í
0x3AB	ŷ ŷ ŷ ŷ ŷ	ŷ ŷ ŷ ŷ ŷ	ŷ ŷ ŷ ŷ ŷ
0x3AC	á á á á á	A A A A A	á á á á á
0x3AD	é é é é é	E E E E E	é é é é é
0x3AE	ń ń ń ń ń	H H H H H	ń ń ń ń ń
0x3AF	í í í í í	I I I I I	í í í í í
x03B0	ú ú ú ú ú ú ú ú ú ú	ŷ ŷ ŷ ŷ ŷ ŷ ŷ ŷ ŷ ŷ	ú ú ú ú ú ú ú ú ú ú
x03B1	α α α α α	A A A A A	α α α α α
x03B2	β β β β β	B B B B B	β β β β β
x03B3	γ γ γ γ γ	Γ Γ Γ Γ Γ	γ γ γ γ γ
x03B4	δ δ δ δ δ	Δ Δ Δ Δ Δ	δ δ δ δ δ
x03B5	ε ε ε ε ε	E E E E E	ε ε ε ε ε
x03B6	ζ ζ ζ ζ ζ	Z Z Z Z Z	ζ ζ ζ ζ ζ
x03B7	η η η η η	H H H H H	η η η η η
x03B8	θ θ θ θ θ	Θ Θ Θ Θ Θ	θ θ θ θ θ
x03B9	í í í í í	I I I I I	í í í í í
x03BA	κ κ κ κ κ	K K K K K	κ κ κ κ κ

x03BB	λ λ λ λ λ	Λ Λ Λ Λ Λ	λ λ λ λ λ
x03BC	μ μ μ μ μ	Μ Μ Μ Μ Μ	μ μ μ μ μ
x03BD	ν ν ν ν ν	Ν Ν Ν Ν Ν	ν ν ν ν ν
x03BE	ξ ξ ξ ξ ξ	Ξ Ξ Ξ Ξ Ξ	ξ ξ ξ ξ ξ
x03BF	ο ο ο ο ο	Ο Ο Ο Ο Ο	ο ο ο ο ο
x03C0	π π π π π	Π Π Π Π Π	π π π π π
x03C1	ρ ρ ρ ρ ρ	Ρ Ρ Ρ Ρ Ρ	ρ ρ ρ ρ ρ
x03C2	ς ς ς ς ς	Σ Σ Σ Σ Σ	ς ς ς ς ς
x03C3	σ σ σ σ σ	Σ Σ Σ Σ Σ'	σ σ σ σ σ
x03C4	τ τ τ τ τ	Τ Τ Τ Τ Τ	τ τ τ τ τ
x03C5	υ υ υ υ υ	Υ Υ Υ Υ Υ	υ υ υ υ υ
x03C6	φ φ φ φ φ	Φ Φ Φ Φ Φ	φ φ φ φ φ
x03C7	χ χ χ χ χ	Χ Χ Χ Χ Χ	χ χ χ χ χ
x03C8	ψ ψ ψ ψ ψ	Ψ Ψ Ψ Ψ Ψ	ψ ψ ψ ψ ψ
x03C9	ω ω ω ω ω	Ω Ω Ω Ω Ω	ω ω ω ω ω
x03CA	Ϊ Ϊ Ϊ Ϊ Ϊ	Ϊ Ϊ Ϊ Ϊ Ϊ	Ϊ Ϊ Ϊ Ϊ Ϊ
x03CB	Ӯ Ӯ Ӯ Ӯ Ӯ	Ӯ Ӯ Ӯ Ӯ Ӯ	Ӯ Ӯ Ӯ Ӯ Ӯ
x03CC	Ӷ Ӷ Ӷ Ӷ Ӷ	Ӷ Ӷ Ӷ Ӷ Ӷ	Ӷ Ӷ Ӷ Ӷ Ӷ
x03CD	ӻ ӻ ӻ ӻ ӻ	ӻ ӻ ӻ ӻ ӻ	ӻ ӻ ӻ ӻ ӻ
x03CE	Ӹ Ӹ Ӹ Ӹ Ӹ	Ӹ Ӹ Ӹ Ӹ Ӹ	Ӹ Ӹ Ӹ Ӹ Ӹ
x03D0	β β	Β Β	β β
x03D1	θ θ	Θ Θ	θ θ
x03D5	φ φ	Φ Φ	φ φ
x03D6	π π	Π Π	π π
x03D8	Ω Ω	Ω Ω	Ω Ω
x03D9	Ϙ Θ	Ϙ Θ	Ϙ Θ
x03DA	Ͳ Ͳ	Ͳ Ͳ	Ͳ Ͳ
x03DB	Ͳ Ͳ	Ͳ Ͳ	Ͳ Ͳ
x03DC	Ͳ Ͳ	Ͳ Ͳ	Ͳ Ͳ
x03DD	Ͳ Ͳ	Ͳ Ͳ	Ͳ Ͳ
x03DE	ʹ ʹ	ʹ ʹ	ʹ ʹ
x03DF	ʹ ʹ	ʹ ʹ	ʹ ʹ
x03E0	ѧ ѧ	ѧ ѧ	ѧ ѧ
x03E1	ܶ ܶ	ܶ ܶ	ܶ ܶ
x03F0	݂ ݂	K K	݂ ݂
x03F1	݂ ݂	P P	݂ ݂
x03F4	݂ ݂	݂ ݂	݂ ݂
x03F5	݂ ݂	E E	݂ ݂

3 Greek extended

x1F00	ጀ ጋ ጋ ጋ ጋ	A A A A A	ጀ ጋ ጋ ጋ ጋ
x1F01	ጀ ጋ ጋ ጋ ጋ	A A A A A	ጀ ጋ ጋ ጋ ጋ
x1F02	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ
x1F03	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ
x1F04	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ
x1F05	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ
x1F06	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ
x1F07	ጀ ጋ ጋ ጋ ጋ ጋ ጋ	A A A A A A A A	ጀ ጋ ጋ ጋ ጋ ጋ ጋ

x1F79	ó ó ó ó ó	O O O O O	ó ó ó ó ó
x1F7A	ù ù ù ù ù	Υ Υ Υ Υ Υ	ù ù ù ù ù
x1F7B	ú ú ú ú ú	Υ Υ Υ Υ Υ	ú ú ú ú ú
x1F7C	ω ω ω ω ω	Ω Ω Ω Ω Ω	ω ω ω ω ω
x1F7D	ώ ώ ώ ώ ώ	Ω Ω Ω Ω Ω	ώ ώ ώ ώ ώ
x1F80	᷇ ᷈ ᷉ ᷊ ᷋	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷋
x1F81	᷇ ᷈ ᷉ ᷊ ᷃	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷃
x1F82	᷇ ᷈ ᷉ ᷊ ᷄	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷄
x1F83	᷇ ᷈ ᷉ ᷊ ᷅	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷅
x1F84	᷇ ᷈ ᷉ ᷊ ᷆	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷆
x1F85	᷇ ᷈ ᷉ ᷊ ᷇	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷇
x1F86	᷇ ᷈ ᷉ ᷊ ᷈	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷈
x1F87	᷇ ᷈ ᷉ ᷊ ᷉	A _I A _I A _I A _I A _I	᷇ ᷈ ᷉ ᷊ ᷉
x1F88	᷈ ᷉ ᷊ ᷋ ᷌	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷋ ᷌
x1F89	᷈ ᷉ ᷊ ᷃ ᷍	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷃ ᷍
x1F8A	᷈ ᷉ ᷊ ᷄ ᷎	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷄ ᷎
x1F8B	᷈ ᷉ ᷊ ᷅ ᷏	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷅ ᷏
x1F8C	᷈ ᷉ ᷊ ᷆ ᷐	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷆ ᷐
x1F8D	᷈ ᷉ ᷊ ᷇ ᷑	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷇ ᷑
x1F8E	᷈ ᷉ ᷊ ᷈ ᷒	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷈ ᷒
x1F8F	᷈ ᷉ ᷊ ᷉ ᷓ	A _I A _I A _I A _I A _I	᷈ ᷉ ᷊ ᷉ ᷓ
x1F90	᷊ ᷋ ᷊ ᷂ ᷌	H _I H _I H _I H _I H _I	᷊ ᷋ ᷊ ᷂ ᷌
x1F91	᷊ ᷃ ᷂ ᷊ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷂ ᷊ ᷌
x1F92	᷊ ᷃ ᷂ ᷄ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷂ ᷄ ᷌
x1F93	᷊ ᷃ ᷂ ᷅ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷂ ᷅ ᷌
x1F94	᷊ ᷃ ᷄ ᷂ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷄ ᷂ ᷌
x1F95	᷊ ᷃ ᷄ ᷃ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷄ ᷃ ᷌
x1F96	᷊ ᷃ ᷅ ᷂ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷂ ᷌
x1F97	᷊ ᷃ ᷅ ᷃ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷃ ᷌
x1F98	᷊ ᷃ ᷄ ᷄ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷄ ᷄ ᷌
x1F99	᷊ ᷃ ᷄ ᷅ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷄ ᷅ ᷌
x1F9A	᷊ ᷃ ᷅ ᷄ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷄ ᷌
x1F9B	᷊ ᷃ ᷅ ᷅ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷅ ᷌
x1F9C	᷊ ᷃ ᷅ ᷆ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷆ ᷌
x1F9D	᷊ ᷃ ᷅ ᷇ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷇ ᷌
x1F9E	᷊ ᷃ ᷅ ᷈ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷈ ᷌
x1F9F	᷊ ᷃ ᷅ ᷉ ᷌	H _I H _I H _I H _I H _I	᷊ ᷃ ᷅ ᷉ ᷌
x1FA0	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊
x1FA1	᷊ ᷊ ᷊ ᷊ ᷋	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷋
x1FA2	᷊ ᷊ ᷊ ᷊ ᷃	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷃
x1FA3	᷊ ᷊ ᷊ ᷊ ᷄	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷄
x1FA4	᷊ ᷊ ᷊ ᷊ ᷅	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷅
x1FA5	᷊ ᷊ ᷊ ᷊ ᷆	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷆
x1FA6	᷊ ᷊ ᷊ ᷊ ᷇	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷇
x1FA7	᷊ ᷊ ᷊ ᷊ ᷈	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷈
x1FA8	᷊ ᷊ ᷊ ᷊ ᷉	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷉
x1FA9	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊
x1FAA	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊
x1FAB	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊
x1FAC	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊
x1FAD	᷊ ᷊ ᷊ ᷊ ᷊	Ω _I Ω _I Ω _I Ω _I Ω _I	᷊ ᷊ ᷊ ᷊ ᷊

x1FE5	đ đ đ đ đ	P P P P P	đ đ đ đ đ
x1FE6	ő ő ő ő ő	ÿ ÿ ÿ ÿ ÿ	ő ő ő ő ő
x1FE7	ű ű ű ű ű ű	ӱ ӱ ӱ ӱ ӱ ӱ	ű ű ű ű ű ű
x1FE8	Ӧ Ӧ Ӧ Ӧ Ӧ	Ӯ Ӯ Ӯ Ӯ Ӯ	Ӧ Ӯ Ӯ Ӯ Ӯ
x1FE9	Ӯ Ӯ Ӯ Ӯ Ӯ	ӯ ӯ ӯ ӯ ӯ	ӯ ӯ ӯ ӯ ӯ
x1FEA	Ӱ Ӱ Ӱ Ӱ Ӱ	Ӳ Ӳ Ӳ Ӳ Ӳ	Ӱ Ӳ Ӳ Ӳ Ӳ
x1FEB	ӱ ӱ ӱ ӱ ӱ	ӳ ӳ ӳ ӳ ӳ	ӱ ӳ ӳ ӳ ӳ
x1FEC	Ӵ Ӵ Ӵ Ӵ Ӵ	ӵ ӵ ӵ ӵ ӵ	Ӵ ӵ ӵ ӵ ӵ
x1FED	ӷ ӷ ӷ ӷ ӷ	Ӷ Ӷ Ӷ Ӷ Ӷ	ӷ Ӷ Ӷ Ӷ Ӷ
x1FEE	Ӹ Ӹ Ӹ Ӹ Ӹ	ӹ ӹ ӹ ӹ ӹ	Ӹ ӹ ӹ ӹ ӹ
x1FEF	ӹ ӹ ӹ ӹ ӹ	ӻ ӻ ӻ ӻ ӻ	ӹ ӻ ӻ ӻ ӻ
x1FE3	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE4	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE5	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE6	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE7	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE8	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FE9	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FEA	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FEB	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FEC	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FED	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ
x1FEE	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ	ጀ ጀ ጀ ጀ ጀ

4 Discussion

4.1 Differences between literal input and LICRs

Under pdfTeX and XeTeX, all input variants select pre-composed characters if possible ([Unicode normalization NFC](#)). Under luatex, accent macros on a literal character are printed using combining diacritical marks. This leads to suboptimal placement of diacritics with many fonts, often resulting in illegible output.

4.1.1 Upcasing

Up to 2022, diacritics were dropped from Greek letters in UPPERCASE in any locale. The new `\MakeUppercase` implementation follows the Unicode standard, where dropping diacritics is a feature of the Greek locale (“el”).

However, “named” accent macros with Greek names behave according to Greek typesetting rules in *all* locales. The current implementation also drops diacritics input via standard accent macros in documents that declare the Greek language with Babel.

Apart from the diacritical marks native to Greek itself, linguists also regularly mark vowel length with macrons (̄) and breves (̄). The length markers are used in dictionaries and textbooks but not part of the Greek orthography. There is no established praxis regarding their handling in uppercase. LaTeX keeps them but the [Unicode reference implementation](#) drops them.

Greek stand-alone diacritics are dropped by `\MakeUppercase` if input via accent macros (e.g. `\accdialytika` or `\'>`).

4.1.2 Sigma, final sigma, and “autosigma”

There are two variants of the lower case letter *sigma*.

The lowering Σ should select a final sigma at the end of a word:
 $\Sigma\Sigma \rightarrow \sigma\varsigma$.

Since 2022, `\MakeLowercase` selects the correct sigma variant for a literal Sigma under Xe/LuaTeX, too.

The `\textSigma` macro is mapped to `\textautosigma` which under 8-bit TeX prints the sigma variant according to its position but, unfortunately, fails with Unicode fonts: $\Sigma\Sigma \rightarrow \sigma\varsigma$.

In LGR, you can add a ZWNJ or use the `\noboundary` macro to prevent the conversion to a final sigma: $\sigma\sigma$.

4.1.3 The one-letter word $\vec{\eta}$ (or)

For disambiguation, the Greek word *or* ($\acute{\eta}$ / $\vec{\eta}$) keeps diacritics in UP-PERCASE:

$\acute{\eta}\,\text{H}\,\vec{\eta}\,\text{H}\,\vec{\eta}\,\text{H} \rightarrow \text{H}\,\text{H}\,\text{H}\,\text{H}\,\text{H}\,\text{H}$.²

However, eta drops diacritics if it is part of a word: $\acute{\eta}\tau\omega\iota$, $\text{H}\tau\omega\iota$, $\varepsilon\pi\tau\alpha\gamma\acute{\eta}$, $\vec{\eta}\tau\omega\iota$, $\text{H}\tau\omega\iota \rightarrow \text{HTOI}$, HTOI , $\text{EPI\kappa\alpha\gamma\eta}$, HTOI , HTOI .³

TODO: The 2022 MakeUppercase handles this for literal input. Currently, this feature is not supported with other input methods.

4.2 Case conversion if the locale is *not* Greek

In order to fix upcasing for Greek, babel-greek maps standard accent macros to “capital” versions.⁴ The default expansion ensures that accents are kept on Latin letters: $\acute{a} \grave{a} \ddot{a} \ddot{a} \rightarrow \acute{A} \grave{A} \ddot{A} \ddot{A}$.⁵

Some characters from the “Greek and Coptic” and “Greek extended” Unicode blocks have no pre-composed uppercase equivalent. Since 2022, upcasing them led to compiling errors under pdfLaTeX. This is fixed in the 2023 LaTeX release by mapping the affected characters to macro representations before case conversions: with pdfLaTeX, diacritics are dropped from literal characters without pre-composed uppercase equivalent also if the text language is *not* Greek, with Xe/LuaTeX, combining characters are used (and may overlap with the base letter).

$\ddot{\epsilon}\ddot{\mu}\ddot{\nu}\ddot{\rho}\ddot{\zeta}\ddot{\eta}\ddot{\pi}\ddot{\iota}\ddot{\nu}\ddot{\rho}\ddot{\zeta}\ddot{\omega}\ddot{\phi} \rightarrow \ddot{\iota}\ddot{\gamma}\ddot{\mu}\ddot{\nu}\ddot{\gamma}\ddot{\alpha}\ddot{\iota}\ddot{\theta}\ddot{\mu}\ddot{\iota}\ddot{\pi}\ddot{\gamma}\ddot{\omega}\ddot{\Omega}\ddot{\Omega}$

4.3 Symbol variants

Several Greek letters have alternative representations which may have a different semantic in a mathematical context. Unicode reserves code points for the alternative representations (naming them “symbol” as opposed to “letter”). Some of these symbols are available as `\var<name>` macros in TeX’s mathematical mode.

In text, they are considered typographical variants without special meaning. The 8-bit LGR font encoding does not allocate code points for

²TODO: The polytonic variant with dasia and oxia used in $\vec{\eta} \dots \vec{\eta}$ (*either* \dots *or*) drops diacritics! By mistake, omission, or intent?

³TODO: It seems the test is only for whitespace on both sides: $\acute{\eta}\,\text{H}\,\vec{\eta}\,\text{H} \rightarrow \text{H}\,\text{H}\,\text{H}\,\text{H}$.

⁴In case of problems, these mappings can be disabled with the `local-MakeUppercase-fixes` language modifier.

⁵If the Language is set to Greek, this works unless the font encoding is 8-bit LGR (where Latin letters are interpreted as Latin transliteration anyway): $\acute{u} \grave{u} \ddot{u} \ddot{u} \rightarrow \Upsilon \Upsilon \breve{\Upsilon} \breve{\Upsilon}$.

them. *Babel-greek* and the *textalpha* package provide the `normalize-symbols` option (ignored with Xe/LuaTeX) that is also used in this test document.

Test the handling of symbol variants in text

literal letter, LICR, `\var<name>`, `\<name>symbol`, literal symbol:

No case change

`βββββ εεεεε θθθθθ ΘΘΘΘΘ ρρρρρ ρρρρρ φφφφφ`

MakeUppercase selects the matching capital letter:

`BBBBB EEEEE ΘΘΘΘΘ ΘΘΘΘΘ KKKKK ΠΠΠΠΠ PPPPP ΦΦΦΦΦ`

MakeLowercase selects the lowercase letter (not the symbol):

`ββ εε θθ θθθ ρρ ππ φφ`

LGR provides two lowercase variants of the archaic letter *stigma*. Both map to GREEK LETTER STIGMA Γ.

The lowercase of Γ is `\textstigma` (not `\textvarstigma`).

4.4 Hiatus

Tonos and psili mark a *hiatus* (break-up) if placed on the first of two vowels that would otherwise form a diphthong. A dialytika must be placed on the second vowel if they are dropped:

`άι άυ έι άυ έι άυ → AI AΥ EI AI AΥ EI AΥ (literal)`
`άι άυ έι άυ έι άυ → AΪ AŶ EΪ AΪ AŶ EΪ AŶ (named+LICR)`
`άι άυ έι άυ έι άυ → AΪ AŶ EΪ AΪ AŶ EΪ AŶ (short+LICR)`
`άι άυ έι άυ έι άυ → AΪ AŶ EΪ AΪ AŶ EΪ AŶ (short+transliteration)`
`άι άυ έι άυ έι άυ → AΪ AŶ EΪ AΪ AŶ EΪ (short+transliteration, reversed)`

The “hiatus feature” does not work with the LGR “input ligatures”. Support for monotonic literal charactes came with the 2022 `\MakeUppercase` implementation changes. Since 2023 it also works with polytonic literals (`άυπνία` → `ΑΥΠΝΙΑ`).

Test the auto-hiatus feature for side-effects:

A B (must keep space after A).

Kerning (compare with source):

`AO AΨ AI AΤ PA OA ΤA ΔΥ [`
`AO AΨ AI AΤ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`
`AO AΨ AΪ AŶ PA OA ΤA ΔΥ [`

Rows 3 … 7: Look-ahead (to check for a hiatus) breaks kerning before ‘A’ with tonos or psili.

Downcasing should keep diacritics (of course, it cannot regenerate “manually” dropped ones): ‘A ũ Ÿ’ → á ü ã.

4.5 Problems with the Latin transliteration

- With monotonic Greek or if `\greekfontencoding` is not LGR, the tilde ~ keeps its default expansion as no-break space (NBSP) and transliterations involving ~ fail. If the `greek` language is loaded with Babel, the tilde accent `\~` as well as combined accents involving the tilde work in all language variants.
- In order to drop accents from input using the Latin transliteration, the characters ', '~, <, and > are suppressed in uppercase. This affects other uses of these characters:

`<`'~^`>` → `<`>` `'\xi'` → `\Xi`

Workaround: use literal typographical quote characters or “french” quotes:

`'\xi'` → `'\Xi'` `«\xi»` → `«\Xi»`

To keep side-effects to a minimum, breathings, combined accents and varia are dropped only with the language variants *polotoniko* or *ancient* as they are not required with monotonic Greek (the current language variant is *polotoniko*).

- The uppercase of spacing composite diacritics including a dialytika `"~" "^"`, appears as apostrophe, as a single " stands for an apostrophe in the LGR transliteration if followed by a space or empty character.
- With the pre-2022 `\MakeUppercase`, the input order is important in combined accents involving the dialytika, e.g., `'"i "``'i` is upcased `\t` → `\I`.