



- [Main Page](#)
- [Help](#)
- [Browse wiki](#)
- [Cookbook](#)
- [Wikijunior](#)
- [Featured books](#)
- [Recent changes](#)
- [Donations](#)
- [Random book](#)
- [Using Wikibooks](#)

#### Community

- [Reading room](#)
- [Community portal](#)
- [Bulletin Board](#)
- [Help out!](#)
- [Policies and guidelines](#)
- [Contact us](#)

#### Tools

- [What links here](#)

[Create account](#) [Log in](#) [Sandbox](#)

Book

[Discussion](#)

Read

[Edit](#)

[View history](#)

Search



Here at Wikibooks we're proud of our **Featured books** - the very best of Wikibooks. Vote for books that you think are worthy of being Featured Books at **Featured book nominations**

# LaTeX/Special Characters

[< LaTeX](#)

In this chapter we will tackle matters related to input encoding, typesetting diacritics and special characters.

In the following document, we will refer to *special characters* for all symbols other than A-Za-z0-9 and English punctuation marks.

This chapter is tightly linked with the font encoding issue. You should have a look at [Fonts](#) on the topic.

Some languages usually need a dedicated input system to ease document writing. This is the case for Arabic, Chinese, Japanese, Korean and others. This specific matter will be tackled in [Internationalization](#).

The rules for producing characters with diacritical marks, such as accents, differ somewhat depending whether you are in text mode, math mode, or the tabbing environment.



## Contents [\[hide\]](#)

### [1 Input encoding](#)

#### [1.1 A technical matter](#)

[Related changes](#)  
[Upload file](#)  
[Special pages](#)  
[Permanent link](#)  
[Page information](#)  
[Cite this page](#)

Languages



Sister projects

[Wikipedia](#)  
[Wikiversity](#)  
[Wiktionary](#)  
[Wikiquote](#)  
[Wikisource](#)  
[Wikinews](#)  
[Wikivoyage](#)  
[Commons](#)  
[Wikidata](#)

Print/export

[Create a collection](#)  
[Download as PDF](#)  
[Printable version](#)

[1.2 Dealing with LaTeX](#)

[1.3 Extending the support](#)

[2 Escaped codes](#)

[3 \*Less than\* < and \*greater than\* >](#)

[4 Euro € currency symbol](#)

[5 Degree symbol for temperature and math](#)

[6 Other symbols](#)

[7 In special environments](#)

[7.1 Math mode](#)

[7.2 Tabbing environment](#)

[8 Unicode keyboard input](#)

[9 External links](#)

[10 Notes and References](#)

## Input encoding [\[edit\]](#)

### A technical matter [\[edit\]](#)

Most of the modern computer systems allow you to input letters of national alphabets directly from the keyboard. If you tried to input these special characters in your LaTeX source file and compiled it, you may have noticed that they do not get printed at all.

A LaTeX source document is a plain text file. A computer stores data in a binary format, that is a sequence of bits (0 and 1). To display a plain text file, we need a code which tells which sequence of bits corresponds to which sequence of characters. This association is called *input encoding*, *character encoding*, or more informally *charset*.

For historical reasons, there are many different input encodings. There is an attempt to unify all

the encoding with a specification that contains all existent symbols that are known from human history. This specification is *Unicode*. It only defines *code points*, which is a number for a symbol, but not the way symbols are represented in binary value. For that, unicode encodings are in charge. There are also several unicode encodings available, UTF-8 being one of them.

The ASCII encoding is an encoding which defines 128 characters on 7 bits. Its widespread use has led the vast majority of encodings to have backward compatibility with ASCII, by defining the first 128 characters the same way. The other characters are added using more bits (8 or more).

This is actually a big issue, since if you do not use the right encoding to display a file, it will show weird characters. What most programs try to do is guess statistically the encoding by analyzing the frequent sequences of bits. Sadly, it is not 100% safe. Some text editors may not bother guessing the encoding and will just use the OS default encoding. You should consider that other people might not be able to display directly your input files on their computer, because the default encoding for text file is different. It does not mean that the user cannot use another encoding, besides the default one, only that it has to be configured. For example, the German umlaut ä on OS/2 is encoded as 132, with Latin1 it is encoded as 228, while in Cyrillic encoding cp1251 this letter does not exist at all. Therefore you should *consider encoding with care*.

The following table shows the default encodings for some operating systems.

Operating system	Default Encodings	
	Western Latin	Cyrillic
Modern Unices (*BSD, Mac OS X, GNU/Linux)	utf-8	utf-8
Mac (before OS X)	applemac	maccyr
Unix (Old)	latin1	koi8-ru

Windows	ansinew, cp1252	cp1251
DOS, OS/2	cp850	cp866nav

UTF-8 and Latin1 are not compatible. It means that if you try to open a Latin1-encoded file using a UTF-8 decoding, it will display odd symbols only if you used accents in it, since both encoding are ASCII superset they encode the *classic* letters the same way. There aren't many advantages in using Latin1 over UTF-8, which is technically superior. UTF-8 is also becoming the most widely used encoding (on the Web, in modern Unices, etc.).



We really urge you to use UTF-8 encoding. It is technically superior to most (all?) encodings, it supports the full Unicode specification (all symbols that ever existed), and is backward compatible with ASCII. Latin1 is not universal, and having multiple encoding around has always been a source of problems.<sup>[1]</sup>

## Dealing with LaTeX [\[edit\]](#)

TeX uses ASCII by default. But 128 characters is not enough to support non-english languages. TeX has its own way to do that with commands for every diacritical marking (see [Escaped codes](#)). But if we want accents and other special characters to appear directly in the source file, we have to tell TeX that we want to use a different encoding.

There are several encodings available to LaTeX:

- ASCII: the default. Only bare english characters are supported in the source file.
- ISO-8859-1 (a.k.a. Latin 1): 8-bits encoding. It supports most characters for latin languages, but that's it.

- UTF-8: a Unicode multi-byte encoding. Supports the complete Unicode specification.
- Others...

In the following we will assume you want to use UTF-8.

There are some *important steps* to specify encoding.

- Make sure your text editor decodes the file in UTF-8.
- Make sure it saves your file in UTF-8. Most text editors do not make the distinction, but some do, such as Notepad++.
- If you are working in a terminal, make sure it is set to support UTF-8 input and output. Some old Unix terminals may not support UTF-8. PuTTY is not set to use UTF-8 by default, you have to configure it.
- Tell LaTeX that the source file is UTF-8 encoded.

```
\usepackage[utf8]{inputenc}
```

`inputenc` [2] package tells LaTeX what the text encoding format of your .tex files is.



If you check the character encoding (e.g. using the Unix `file` command), be sure that your file contains at least one special character, otherwise it will be recognized as ASCII (which is logical since UTF-8 is a superset of ASCII).

The `inputenc` package allows as well the user to change the encoding *within the document* by means of the command `\inputencoding{'encoding name'}`.

```

\usepackage[utf8]{inputenc}
% ...
% In this area
% The UTF-8 encoding is specified.
% ...
\inputencoding{latin1}
% ...
% Here the text encoding is specified as ISO Latin-1.
% ...
\inputencoding{utf8}
% Back to the UTF-8 encoding.
% ...

```

## Extending the support [\[edit\]](#)

The LaTeX support of UTF-8 is fairly specific: it includes only a limited range of unicode input characters. It only defines those symbols that are known to be available with the current *font encoding*. You might encounter a situation where using UTF-8 might result in error:

```
! Package inputenc Error: Unicode char \u8:ũ not set up for use with LaTeX.
```

This is due to the utf8 definition not necessarily having a mapping of all the character glyphs you are able to enter on your keyboard. Such characters are for example

ŷ Ŷ ù Û ë Ë ĩ Ĭ

In such case, you may try need to use the `utf8x` option to define more character combinations. `utf8x` is not officially supported, but can be viable in some cases. However it might break up compatibility with some packages like `csquotes`.

Another possibility is to stick with `utf8` and to define the characters yourself. This is easy:

```
\DeclareUnicodeCharacter{'codepoint'}{'TeX sequence'}
```

where `codepoint` is the unicode codepoint of the desired character. `TeX sequence` is what to print when the character matching the codepoint is met. You may find codepoints on this [site](#).

Codepoints are easy to find on the web. Example:

```
\DeclareUnicodeCharacter{0177}{\^y}
```

Now inputting `'y'` will effectively print `'y'`.

With XeTeX and LuaTeX the inputenc package is no longer needed. Both engines support UTF-8 directly and allow the use of TTF and OpenType fonts to support Unicode characters. See the [Fonts](#) section for more information.

## Escaped codes [\[edit\]](#)

In addition to direct UTF-8 input, LaTeX supports the composition of special characters. This is convenient if your keyboard lacks some desired accents and other diacritics.

The following accents may be placed on letters. Although 'o' letter is used in most of the examples, the accents may be placed on any letter. Accents may even be placed above a "missing" letter; for example, `\~{ }` produces a tilde over a blank space.

The following commands may be used only in paragraph (default) or LR (left-right) mode.

LaTeX command	Sample	Description
<code>\`{o}</code>	ò	grave accent
<code>\'{o}</code>	ó	acute accent
<code>\^{o}</code>	ô	circumflex
<code>\ \"{o}</code>	ö	umlaut, trema or dieresis
<code>\H{o}</code>	ő	long Hungarian umlaut (double acute)
<code>\~{o}</code>	õ	tilde
<code>\c{c}</code>	ç	cedilla
<code>\k{a}</code>	ą	ogonek
<code>\l</code>	ł	barred l (l with stroke)
<code>\={o}</code>	ō	macron accent (a bar over the letter)
<code>\b{o}</code>	ȕ	bar under the letter
<code>\.{o}</code>	ȏ	dot over the letter
<code>\d{u}</code>	ș	dot under the letter
<code>\r{a}</code>	å	ring over the letter (for å there is also the special command <code>\aa</code> )
<code>\u{o}</code>	ǫ	breve over the letter
<code>\v{s}</code>	š	caron/háček ("v") over the letter



<code>\t{oo}</code>	oö	"tie" (inverted u) over the two letters
<code>\o</code>	ø	slashed o (o with stroke)

To place a diacritic on top of an i or a j, its dot has to be removed. The dotless version of these letters is accomplished by typing `\i` and `\j`. For example:

- `\^{\i}` should be used for i circumflex 'î';
- `\"\i` should be used for i umlaut 'ï'.

If a document is to be written completely in a language that requires particular diacritics several times, then using the right configuration allows those characters to be written directly in the document. For example, to achieve easier coding of umlauts, the babel package can be configured as `\usepackage[german]{babel}`. This provides the short hand "o for `\o`. This is very useful if one needs to use some text accents in a label, since no backslash will be accepted otherwise.

More information regarding language configuration can be found in the [Internationalization](#) section.

## *Less than < and greater than >* [\[edit\]](#)

The two symbols '<' and '>' are actually ASCII characters, but you may have noticed that they will print 'ı' and 'ç' respectively. This is a font encoding issue. If you want them to print their real symbol, you will have to use another font encoding such as T1, loaded with the `fontenc` package. See [Fonts](#) for more details on font encoding.

Alternatively, they can be printed with dedicated commands:

`\textless`

`\textgreater`

## Euro € currency symbol [\[edit\]](#)

---

When writing about money these days, you need the [euro sign](#). The `textcomp` package features a `\texteuro` command which gives you the euro symbol as supplied by your current text font.

Depending on your chosen font this may be quite far from the official symbol.

An official version of the euro symbol is provided by [eurosym](#). Load it in the preamble (optionally with the `official` option):

```
\usepackage[official]{eurosym}
```

then you can insert it with the `\euro{}` command. Finally, if you want a euro symbol that matches with the current font style (e.g., bold, italics, etc.) you can use a different option:

```
\usepackage[gen]{eurosym}
```

again you can insert the euro symbol with `\euro{}`.

Alternatively you can use the [marvosym](#) package which also provides the official euro symbol.

```
\usepackage{marvosym}
```

```
% ...
```

`\EUR{}`

Now that you have succeeded in printing a euro sign, you may want the '€' on your keyboard to actually print the euro sign as above. There is a simple method to do that. You must make sure you are using UTF-8 encoding along with a working `\euro{}` or `\EUR{}` command.

```
\DeclareUnicodeCharacter{20AC}{\euro{}}
```

*% or*

```
\DeclareUnicodeCharacter{20AC}{\EUR{}}
```

Complete example:

```
\usepackage[utf8]{inputenc}
```

```
\usepackage{marvosym}
```

```
\DeclareUnicodeCharacter{20AC}{\EUR{}}
```

## Degree symbol for temperature and math [\[edit\]](#)

A common mistake is to use the `\circ` command. It will not print the correct character. Use the `textcomp` package instead, which provides a `\textdegree` command.

```
\usepackage{textcomp}
```

*%...*

A `\textdegree` angle.

For temperature, you can use the same command or opt for the `gensymb` package and write

```
\usepackage{gensymb}
```

```
%...
```

```
17\degree~C
```

or

```
17\celsius
```

Some keyboard layouts feature the degree symbol, you can use it directly if you are using UTF-8 and `textcomp`. For better results (font quality) we recommend the use of an appropriate font, like `lmodern`:

```
\usepackage[utf8]{inputenc}
```

```
\usepackage{lmodern}
```

```
\usepackage{textcomp}
```

```
% ...
```

```
17°~C
```

# Other symbols [\[edit\]](#)

LaTeX has *lots* of symbols at its disposal. The majority of them are within the mathematical domain, and later chapters will cover how to get access to them. For the more common text symbols, use the following commands:

Command	Symbol	Command	Symbol
<code>\%</code>	<code>%</code>	<code>\#</code>	<code>#</code>
<code>\\$</code>	<code>\$</code>	<code>\&amp;</code>	<code>&amp;</code>
<code>\{</code>	<code>{</code>	<code>\}</code>	<code>}</code>
<code>\_</code>	<code>-</code>	<code>\S</code>	<code>§</code>
<code>\P</code>	<code>¶</code>	<code>\dag</code>	<code>†</code>
<code>\ddag</code>	<code>‡</code>	<code>\textbackslash</code>	<code>\</code>
<code>\textbar</code>	<code> </code>	<code>\textless</code>	<code>&lt;</code>
<code>\textgreater</code>	<code>&gt;</code>	<code>\textemdash</code>	<code>—</code>
<code>\textendash</code>	<code>-</code>	<code>\textregistered</code>	<code>®</code>
<code>\texttrademark</code>	<code>™</code>	<code>\textquestiondown</code>	<code>¿</code>
<code>\textexclamdown</code>	<code>¡</code>	<code>\textcircled{a}</code>	<code>Ⓐ</code>
<code>\textsuperscript{a}</code>	<code><sup>a</sup></code>	<code>\copyright</code>	<code>©</code>
<code>\pounds</code>	<code>£</code>		

Not mentioned in above table, tilde (~) is used in LaTeX code to produce [non-breakable space](#). To get printed tilde sign, either write `\~{ }` or `\textasciitilde{ }`. And a visible space  can be created with `\textvisiblespace`.

For some more interesting symbols, the Postscript ZipfDingbats font is available thanks to the [pifont](#) package. Add the declaration to your preamble: `\usepackage{pifont}`. Next, the command `\ding{number}`, will print the specified symbol. Here is a table of the available symbols:

32		33	✂	34	✂	35	✂	36	✂	37	✂	38	⌚	39	⌚
40	✈	41	✉	42	✉	43	✉	44	✉	45	✉	46	✉	47	✉
48	✉	49	✉	50	✉	51	✓	52	✓	53	✕	54	✕	55	✕
56	✕	57	✕	58	✕	59	✕	60	✕	61	✕	62	✕	63	✕
64	✕	65	✕	66	✕	67	✕	68	✕	69	✕	70	✕	71	✕
72	★	73	☆	74	★	75	☆	76	★	77	★	78	★	79	★
80	☆	81	✱	82	✱	83	✱	84	✱	85	✱	86	✱	87	✱
88	✱	89	✱	90	✱	91	✱	92	✱	93	✱	94	✱	95	✱
96	✱	97	✱	98	✱	99	✱	100	✱	101	✱	102	✱	103	✱
104	✱	105	✱	106	✱	107	✱	108	●	109	○	110	■	111	□
112	□	113	□	114	□	115	▲	116	▼	117	◆	118	◆	119	◆
120		121		122		123	•	124	•	125	•	126	•		
		161	♩	162	♩	163	♩	164	♥	165	♠	166	♣	167	♣
168	♣	169	♦	170	♥	171	♠	172	①	173	②	174	③	175	④
176	⑤	177	⑥	178	⑦	179	⑧	180	⑨	181	⑩	182	①	183	②
184	③	185	④	186	⑤	187	⑥	188	⑦	189	⑧	190	⑨	191	⑩
192	①	193	②	194	③	195	④	196	⑤	197	⑥	198	⑦	199	⑧
200	⑨	201	⑩	202	①	203	②	204	③	205	④	206	⑤	207	⑥
208	⑦	209	⑧	210	⑨	211	⑩	212	→	213	→	214	↔	215	↕
216	↖	217	→	218	↗	219	→	220	→	221	→	222	→	223	→
224	→	225	→	226	→	227	→	228	→	229	→	230	→	231	→
232	→	233	→	234	→	235	→	236	→	237	→	238	→	239	→
		241	→	242	→	243	→	244	→	245	→	246	→	247	→
248	→	249	→	250	→	251	→	252	→	253	→	254	→		

In special environments [\[edit\]](#)

**Math mode** [\[edit\]](#)

Several of the above and some similar accents can also be produced in math mode. The following

commands may be used only in math mode.

LaTeX command	Sample	Description	Text-mode equivalence
<code>\hat{o}</code>	$\hat{o}$	circumflex	<code>\^</code>
<code>\widehat{oo}</code>	$\widehat{oo}$	wide version of <code>\hat</code> over several letters	
<code>\check{o}</code>	$\check{o}$	vee or check	<code>\v</code>
<code>\tilde{o}</code>	$\tilde{o}$	tilde	<code>\~</code>
<code>\widetilde{oo}</code>	$\widetilde{oo}$	wide version of <code>\tilde</code> over several letters	
<code>\acute{o}</code>	$\acute{o}$	acute accent	<code>\'</code>
<code>\grave{o}</code>	$\grave{o}$	grave accent	<code>\`</code>
<code>\dot{o}</code>	$\dot{o}$	dot over the letter	<code>\.</code>
<code>\ddot{o}</code>	$\ddot{o}$	two dots over the letter (umlaut in text-mode)	<code>\"</code>
<code>\breve{o}</code>	$\breve{o}$	breve	<code>\u</code>
<code>\bar{o}</code>	$\bar{o}$	macron	<code>\=</code>
<code>\vec{o}</code>	$\vec{o}$	vector (arrow) over the letter	

When applying accents to letters *i* and *j*, you can use `\imath` and `\jmath` to keep the dots from interfering with the accents:

LaTeX command	Sample	Description	Sample with upper dot
<code>\hat{\imath}</code>	$\hat{\imath}$	circumflex on letter <i>i</i> without upper dot	$\hat{\dot{\imath}}$

`\vec{\jmath}`



vector (arrow) on letter j without upper dot  $\vec{j}$

## Tabbing environment [\[edit\]](#)

Some of the accent marks used in running text have other uses in the tabbing environment. In that case they can be created with the following command:

- `\a'` for an acute accent
- `\a`` for a grave accent
- `\a=` for a macron accent

## Unicode keyboard input [\[edit\]](#)

Some operating systems provide a keyboard combination to input any Unicode code point, the so-called *unicode compose key*.

Many X applications (\*BSD and GNU/Linux) support the Ctrl+Shift+u combination. A 'u' symbol should appear.

Type the code point and press enter or space to actually print the character. Example:

```
<Ctrl+Shift+u> 20AC <space>
```

will print the euro character.

Desktop environments like GNOME and KDE may feature a customizable compose key for more memorable sequences.

Xorg features advanced keyboard layouts with variants that let you enter a lot of characters easily






Wikipedia has related information at **Unicode input**





with combination using the appropriate modifier, like Alt Gr. It highly depends on the selected layout+variant, so we suggest you to play a bit with your keyboard, preceeding every key and dead key with the Alt Gr modifier.

## External links [\[edit\]](#)

- [A few other LaTeX accents and symbols](#) 
- [NASA GISS: Accents](#) 
- [PDF document with a lengthy list of symbols provided by various packages](#) 

## Notes and References [\[edit\]](#)

1. [↑](#) For a quick explanation on character sets, see [this article on Joel Spolski's blog](#) .
2. [↑](#) For a detailed information on the package, see [complete specifications written by the package's authors](#) .

[Previous: List Structures](#)

[Index](#)

[Next: Internationalization](#)

[Category: LaTeX](#)

This page was last modified on 24 September 2014, at 11:37.

Text is available under the [Creative Commons Attribution-ShareAlike License](#).; additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#).

[Privacy policy](#) [About Wikibooks](#) [Disclaimers](#) [Developers](#) [Mobile view](#)

