Typesetting Greek Texts with GreekTeX

K J Dryllerakis

Imperial College
Department of Computing
London
(kd@doc.ic.ac.uk)

This document describes GreekTeX, yet another something-TeX application, a package for typesetting greek texts. The following pages are intended to be a quick tutorial on the package and not a detailed account of the package. It is assumed that the GreekTeX package is already installed at you site and both the greek format file (`greek.fmt`) and the alternative GreekTeX macros (`greektex.tex`) are accesible to the user. We also assume that all the recomended `kd` fonts exist in the font path. If this is not the case please consult the installation procedure document which gives all the information needed to install the package in your site. In the discussion to follow, more TeXnical material is presented in a finer print.

The Greek Mode And Fonts

In order to typeset greek text, you have to advise TeX to enter a "Greek Mode"[1]. Entering the greek mode is achieved by typing `\begingreek` at the position were we want the greek text to begin. The greek mode will be active until an `\endgreek` is typed, when we return to the previous state. Inside the greek mode, the latin characters typed correspond to the greek alphabet and diacritic marcs. The family changing commands (`\sl`, `\bf`, `\tt`, `\it`) are automatically switched to correspond to the proper greek font families and the classical greek font corresponds to the `\gr` or `\rg` command. Roman letters are still available through the `\rm` and `\l`*oldsequence* commands (e.g. `\lsl`). When we re-enter the greek mode, the font used last is remembered and typesetting continues from the state we left of in the previous greek mode.

TeXnically, entering the greek mode means *i.* to redifine the `\catcode` values for several characters (mostly accents and breathing symbols), *ii.* to declare that we are currently typesetting in greek and use the correct hyphenation tables, *iii.* switch to the point size and font used last, and *iv.* redifine the control sequences corresponding to font family switching (e.g. `\sl` etc.).

The font families that are currently avaible are all based on the original designs by Sylvio Levi and Haralambous. The families (and the corresponding font switching commands) are: classical greek (`\gr` or `\rg`), slanted greek (`\sl`), greek italics (`\it`), boldface greek (`\bf`) and typewriter greek (`\tt`). Remember that these families are only available inside greek mode.

A set of point-setting commands are made available from the greek format. Three sizes are prespesified for ten, nine, and eight points respectevly. The corresponding commands are `\grtenpoint`, `\grninepoint`, and `\greightpoint`. Since not all fonts exist in the sizes need, scaled fonts are used instead. Remeber that re-entry in the greek mode means picking up from exactly where we previously left i.e. from the same point size. If point-size changes are to remain local the user is encouraged to enclose them in groups.

The real names for the control sequences switching to greek font families are `\gr`, `\git`, `\gsl`, `\gbf`, `gtt`. These commands can be used even outside greek mode to typeset greek short texts but it is not recomended. Note that the LaTeX version of GreekTeX (the greek style file `greek.sty`) uses only these font changing commands.

Since the commands `\begingreek` and `\endgreek` can be quite tedious to repeat, they can be both substituted either by the character `$` or `|` if at the beginning of your file you specify `\greekdelims{dollar}` or `\greekdelims{bar}` repsectively. In case the `$` character is used, math mode is accessed by the control sequence `\math`.

The Greek Alphabet

Since TeX was designed to help portability of files, most implemetations support only 128 character input (file transfer is also based on the 128 ascii set). That means we have to use the Latin alphanumeric characters

[1] The idea of a "greek environment" was first introduced by Sylvio Levi in his greek fonts; the top level macros of our package are inherited from this first approach.

to represent the greek ones. The following scheme is used to represent greek characters inside the greek mode:

α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο	π	ρ	ς	τ	υ	φ	χ	ψ	ω	ς
a	b	g	d	e	z	h	j	i	k	l	m	n	x	o	p	r	s	t	u	f	q	y	w	c

It is a common practise in Greece to substitue the characters corresponding to the higher ascii codes by greek characters so that users are able to see the greek characters directly on the screen. It is fairly easy to write a program to transform any such ascii output to the corresponding “latin” representation of the greek characters and feed it directly to TEX’s mouth. For example such a translation program is supplied with the current distribution of the emTEX package for DOS based machines.

Accents, Breathing And Punctuation

Three different accents are encountered in greek texts. In order to accent a vowel simply type `'`, `` ` ``, or `~` to get an acute, grave or circumflex accent repsectively. Breathing signs are achieved in the same way by preceding the vowel (and any accent that it may have) with `<` for rough and `>` for smooth breathing. Breathings can also be used before a greek rho (ρ). Iota subscript is achieved by typing `|` *after* vowel. The diairesis sign is achieved by presiding a (maybe accented) vowel by `"`. Greek punctutation marcs are achieved according to the following table:

.	,	·	:	!	;	’	«	»
.	,	;	:	!	?	’’	((	))

For example, concider the following quotation from Xenophon :

> Ἐν οὐδεμία γάρ πόλει τὸ βὲλτιστον εὔνουν ἐστὶ τῷ δῆμῳ, ἀλλὰ τὸ κάκιστον ἐν ἑκάστῃ ἐστὶ πόλει εὔνουν τῷ δῆμῳ· οἱ γάρ ὅμοιοι τοῖς ὁμοίοις εὔνοι εἰσι.

was typeset by

```
>En o>udem'ia g'ar p'olei t`o b`eltiston e>'unoun >est`i t~w|
d~hmw|, >all`a t`o k'akiston >en <ek'asth| >est`i
p'olei e>'unoun t~w| d~hmw|; o<i g'ar <'omoioi to~ic
<omo'ioic e>'unoi e>isi.
```

Note the use of punctuation and breathing marks before the vowels.

All accents and breathings are recognised as normal letters insode the greek mode. They will print for themselves when they stand alone but form ligatures when preciding a vowel. The ligatures do all the work for us; the kerning between accents or breathing and capital vowels is also done through kerning. This simplifies the macros used and the greek character set is usable with only a few `\catcode` definitions. All the rest of the mechanisms supplied in the greek format are only to facilitate the usage of the fonts in standard text and macro creation.

Hyphenation

If you are working with the greek format file, then correct hyphenation is assured for greek text. The hyphenation patterns supplied in the file are the ones compiled by Haralambous in France. If you are working with the file `greektex.tex` then no hyphenation patterns are loaded; this fact makes the macros of `greektex.tex` suitable for only small greek texts. GREEKTEX, takes full advantage of the multilingual characteristics of TEX version 3.0 or higher by declaring a new language and identifying the correct patterns for this new language. The result is correct hyphenation for texts containing both greek and latin texts.

In case you want to use GREEKTEXwith another multilingual package, it is important to find out the limitation of your TEXversion. It is possible that the number of allowed hyphenation patterns exceeds the number needed. Remember that plain TEX uses 4447 patterns and GREEKTEX needs a futher 1170.

It is possible to see the proposed hyphenation for a word using a macro similar to `\showhyphens` (supplied with plain TEX). The macro `\showgreekhyphens` can be used succesfully only inside greek mode. This is due to the fact that the argument for the macro has to be scanned with the correct `\catcode` values active.

Building Macros

Let us now turn to some examples that will demonstrate the use of the GreekTeX macros within your own definitions.

ΠΑΛΑΙΟΘΕΝ ΕΛΛΗΝΙΣ

Καυχιέται ἡ ᾿Αντιόχεια	γιὰ τὰ λαμπρά της κτίρια,
καὶ τούς ὡραίους της δρόμους·	γιὰ τὴν περὶ αὐτήν
θαυμάσιαν ἐξοχήν,	καὶ γιὰ τὸ μέγα πλῆθος
τῶν ἐν αὐτῇ κατοίκων.	Καυχιέται ποὺ εἶν' ἡ ἕδρα
ἐνδόξων βασιλέων·	καὶ γιὰ τοὺς καλλιτέχνας
καὶ τοὺς σοφούς ποὺ ἔχει,	καὶ γιὰ τοὺς βαθυπλούτους
καὶ γνωστικοὺς ἐμπόρους.	Μὰ πιὸ πολὺ ἀσυγκρίτως
ἀπ' ὅλα, ἡ ᾿Αντιόχεια	καυχιέται ποὺ εἶναι πόλις
παλαιόθεν ἑλληνίς·	τοῦ ῎Αργους συγγενής:
ἀπ' τὴν ᾿Ιώνη ποὺ	ἱδρύθη ὑπὸ ᾿Αργείων
ἀποίκων πρὸς τιμὴν	τῆς κόρης τοῦ ᾿Ινάχου.

Κ. Καβάφης

Let us try to typeset this poem written by a greek poet of the 20th century. The peculiar format is the original one required by the poet. It is immediatly apparent that a command taking two arguments can be used to typeset each verse; the arguments will be the left and right part of each verse. The only important problem encountered is the following: when TeX is reading the arguments of a command, it immedietly assigns `\catcode` values to the tokens read. So, if the correct `\catcode` values are not active at the time the arguments are scanned, TeX is much confused. Consequently, macros that take as arguments greek text must be called from within greek mode. Having that in mind we can define a macro `\verse` that will correctly typeset one verse of the poem. The macro is defined as follows:

```
\def\verse#1#2{\line{#1\hfill#2}}
```

It is assumed of course that the `\hsize` required for the document is already set. The rest is easy.

```
$
\centerline{\bf PALAIOJEN ELLHNIS}
\smallskip
\verse{Kauqi'etai <h >Anti'oqeia}{gi`a t`a lampr'a thc kt'iria,}
                                   ⋮
\verse{>apo'ikwn pr`oc tim`hn}{t~hc k'orhc to~u >In'aqou.}
\bigskip
\leftline{K. Kab'afhc}
$
```

7.2 Τό σύστημα Κέντρου Μάζας 172

7.2 Το σύστημα κέντρου Μάζας

Είναι συχνά βολικό να περιγράφεται η κίνηση του συστήματος σε πλαίσιο αναφοράς στο οποίο το κέντρο μάζας ηρεμεί στην αρχή των αξόνων. (Σε βαρυτικό πεδίο, το σύστημα αυτό είναι επιταγχυνόμενο, μη αδρανειακό, παρόλα αυτά είναι πάλι χρήσιμο.) Καλείται το σύστημα *κέντρου μάζας* (ΚΜ). Θα συμβολίζουμε τα μεγέθη που ανδαφέρονται σ'αυτό με έναν αστερίσκο.

Η σχετική θέση r είναι βέβαια ανεξάρτητη από το πώς επιλέγεται η αρχή, έτσι ώστε θέτοντας $R^* = 0$ στήν (7.7) βρίσκουμε

$$\mathbf{r}_1^* = \frac{m_2}{M}\mathbf{r}, \qquad \mathbf{r}_2^* = -\frac{m_1}{M}\mathbf{r} \tag{7.12}$$

Στό σύστημα αυτό, οι ορμές των δύο σωματίων είναι (ίσες και) αντίθετες,

$$m_i r_i^* = -m_2 r_2^* = \mu r = p^* \tag{7.13}$$

Όπως θα δούμε καθαρά αργότερα, είναι συχνά βολικό να επιλύουμε ενα πρόβλημα πρώτα στο σύστημα ΚΜ.

⋮

Let us now turn to our second example which will demonstrate scientific typesetting in greek. The example is taken from the book *Classical Mechanics* by T.W.B. Kibble in its greek translation. Let us suppose that we are trying to typeset this book; how would we go about to build the macros needed? First of all we will be using a lot of math formulae and therefore we are in desperate need of the start and end math symbol. Since we are using modern greek and no iota ligatures are needed, we can redifine the bar `|` to stand for the begin and end greek by `\greekdelims{bar}`. This might turn out to be convenient if we want to include latin words in the text. Now for the macros. Starting with the headline appearing at the top of the page we need to define the

`\headline` macro. This must be done within greek mode to correctly interprent the tokens scanned:

```
\begingreek
\global\headline={\hfil {\git T'o s'usthma K'entrou M'azac}\ \folio}
\endgreek
```

It might be wiser to change the `\git` command either to `\grtenpoint\git` or `\tengri` to ensure that the correct font is selected when the output routine is called.

To typeset math formulae you only have to use the standard TEX rules. Even within greek mode entering the math mode has exactly the same effect as from the plain TEX mode. This is not entirely true since the font changing sequences are redefined. If you want boldface characters within the math mode you have to use the `\lbf` command and the same for any other family. See how is it really is:

```
\hsize=12truecm
\begingreek
\grtenpoint % Switch to ten point characters
\gr % Default font
\noindent {\bf 7.2 To s'usthma k'entrou M'azac}
\medskip
E'inai suqn'a bolik'o na perigr'afetai h k'inhsh tou sust'hmatoc se
pla'isio anafor'ac sto opo'io to k'entro m'azac hreme'i sthn arq'h twn
ax'onwn. (Se barutik'o ped'io, to s'usthma aut'o e'inai epitagqun'omeno, mh
adraneiak'o, par'ola aut'a e'inai p'ali qr'hsimo.) Kale'itai to s'usthma
{\it k'entrou m'azac} (KM). Ja sumbol'izoume ta meg'ejh pou andaf'erontai
s''aut'o me 'enan aster'isko.
H sqetik'h j'esh $r$  e'inai b'ebaia anex'arthth ap'o to p'wc
epil'egetai h arq'h, 'etsi 'wste j'etontac $R^*=0$  st'hn
(7.7) br'iskoume
$${\lbf r}^*_1={m_2\over M}{\lbf r},
       \qquad {\lbf r}^*_2=-{m_1\over M}{\lbf r}\eqno(7.12)$$
St'o s'usthma aut'o, oi orm'ec twn d'uo swmat'iwn e'inai
('isec kai) ant'ijetec,
$$m_ir^*_i = -m_2r^*_2=\mu r = p^*\eqno(7.13)$$
'Opwc ja do'ume kajar'a arg'otera, e'inai suqn'a bolik'o
na epil'uoume ena pr'oblhma pr'wta sto s'usthma KM.
\endgreek
```

Final Remarks

Except from the macros explained above, the greek format also has a set ofgeneral use macros. It is possinble that in certain cases, greek mode will be the dominant mode. To switch to a permenant greek mode use the command `\setgreek`. If you want to restore the original TEX mode (including restoring any reassignents of `$` or `|`) use the command `\setlatin`.

These commands are based on the internal commands `\gr@@km@de` and `\l@tinmode`. Both define the `\catcode`s for the mode. A macro-builder can use them to create different environments e.g. a command to switch to english from a global greek mode. Just remeber that you will be on your own so you will have to switch to the enclish language manually by a command `\language0`.