

The `setouterhbox` package

Heiko Oberdiek
<oberdiek@uni-freiburg.de>

2006/08/26 v1.4

Abstract

If math stuff is set in an `\hbox`, then TeX performs some optimization and omits the implicate penalties `\binoppenalty` and `\relpenalty`. This packages tries to put stuff into an `\hbox` without getting lost of those penalties.

Contents

1	Documentation	1
1.1	Introduction	1
1.2	Acknowledgement	2
1.3	Usage	2
1.4	Option <code>hyperref</code>	3
1.5	Example	3
2	Implementation	3
2.1	Package start stuff	3
2.2	Main part	4
2.3	Environment support	7
2.4	Option <code>hyperref</code>	7
3	Installation	8
3.1	Some details for the interested	8
4	References	9
5	History	9
	[2005/10/05 v1.0]	9
	[2005/10/07 v1.1]	9
	[2005/10/18 v1.2]	9
	[2006/02/12 v1.3]	9
	[2006/08/26 v1.4]	9
6	Index	10

1 Documentation

1.1 Introduction

There is a situation in `hyperref`'s driver for dvips where the user wants to have links that can be broken across lines. However dvips doesn't support the feature. With option `breaklinks` `hyperref` sets the links as usual, put them in a box and write the link data with box dimensions into the appropriate `\specials`. Then, however, it does not set the complete unbreakable box, but it unwrappes the material inside to allow line breaks. Of course line breaking and glue setting will falsify the link dimensions, but line breaking was more important for the user.

1.2 Acknowledgement

Jonathan Fine, Donald Arsenau and me discussed the problem in the newsgroup `comp.text.tex` where Damian Menscher has started the thread, see [1].

The discussion was productive and generated many ideas and code examples. In order to have a more permanent result I wrote this package and tried to implement most of the ideas, a kind of summary of the discussion. Thus I want and have to thank Jonathan Fine and Donald Arsenau very much.

Two weeks later David Kastrup (posting in `comp.text.tex`, [2]) remembered an old article of Michael Downes ([3]) in TUGboat, where Michael Downes already presented the method we discuss here. Nowadays we have ε -TeX that extends the tool set of a TeX macro programmer. Especially useful ε -TeX was in this package for detecting and dealing with erroneous situations.

However also nowadays a perfect solution for the problem is still missing at macro level. Probably someone has to go deep in the internals of the TeX compiler to implement a switch that let penalties stay where otherwise TeX would remove them for optimization reasons.

1.3 Usage

Package loading. L^ATeX: as usually:

```
\usepackage{setouterhbox}
```

The package can also be included directly, thus plain-TeX users write:

```
\input setouterhbox.sty
```

Register allocation. The material will be put into a box, thus we need to know these box number. If you need to allocate a new box register:

L^ATeX: `\newsavebox{\langle name \rangle}`

plain-TeX: `\newbox\langle name \rangle`

Then `\langle name \rangle` is a command that held the box number.

Box wrapping. L^ATeX users put the material in the box with an environment similar to `lrbox`. The environment `setouterhbox` uses the same syntax and offers the same features, such as verbatim stuff inside:

```
\begin{setouterhbox}{\langle box number \rangle}... \end{setouterhbox}
```

Users with plain-TeX do not have environments, they use instead:

```
setouterhbox{\langle box number \rangle}... \endsetouterhbox
```

In both cases the material is put into an `\hbox` and assigned to the given box, denoted by `\langle box number \rangle`. Note the assignment is local, the same way `lrbox` behaves.

Unwrapping. The box material is ready for unwrapping:

```
\unhbox\langle box number \rangle
```

1.4 Option hyperref

Package `url` uses math mode for typesetting urls. Break points are inserted by `\binoppenalty` and `\relpenalty`. Unhappily these break points are removed, if `hyperref` is used with option `breaklinks` and drivers that depend on `pdfmark:dvips`, `vtexpdfmark`, `textures`, and `dvipsone`. Thus the option `hyperref` enables the method of this package to avoid the removal of `\relpenalty` and `\binoppenalty`. Thus you get more break points. However, the link areas are still wrong for these drivers, because they are not supporting broken links.

Note, you need version 2006/08/16 v6.75c of package `hyperref`, because starting with this version the necessary hook is provided that package `setouterhbox` uses.

```
\usepackage[...]{hyperref}[2006/08/16]
\usepackage[hyperref]{setouterhbox}
```

Package order does not matter.

1.5 Example

```
1 (*example)
2 \documentclass[a5paper]{article}
3 \usepackage{url}[2005/06/27]
4 \usepackage{setouterhbox}
5
6 \newsavebox{\testbox}
7
8 \setlength{\parindent}{0pt}
9 \setlength{\parskip}{2em}
10
11 \begin{document}
12 \raggedright
13
14 \url{http://this.is.a.very.long.host.name/followed/%
15 by/a/very_long_long_long_path.html}%
16
17 \sbox\testbox{%
18   \url{http://this.is.a.very.long.host.name/followed/%
19   by/a/very_long_long_long_path.html}%
20 }%
21 \unhbox\testbox
22
23 \begin{setouterhbox}{\testbox}%
24   \url{http://this.is.a.very.long.host.name/followed/%
25   by/a/very_long_long_long_path.html}%
26 \end{setouterhbox}
27 \unhbox\testbox
28
29 \end{document}
30 </example>
```

2 Implementation

Internal macros are prefixed by `\setouterhbox`, `@` is not used inside names, thus we do not need to care of its catcode if we are not using it as \LaTeX package.

2.1 Package start stuff

```
31 (*package)
32
33 Prevent reloading more than one, necessary for plain- $\text{\TeX}$ :
34 \begingroup
35   \expandafter\let\expandafter\x\csname ver@setouterhbox.sty\endcsname
36 \ifcase 0%
```

```

35 \ifx\x\relax % plain
36 \else
37 \ifx\x\empty % LaTeX
38 \else
39 1%
40 \fi
41 \fi
42 \else
43 \expandafter\ifx\csname PackageInfo\endcsname\relax
44 \def\x#1#2{%
45 \immediate\write-1{Package #1 Info: #2.}%
46 }%
47 \else
48 \def\x#1#2{\PackageInfo{#1}{#2, stopped}}%
49 \fi
50 \x{setouterhbox}{The package is already loaded}%
51 \endgroup
52 \expandafter\endinput
53 \fi
54 \endgroup
Package identification:
55 \begingroup
56 \expandafter\ifx\csname ProvidesPackage\endcsname\relax
57 \def\x#1#2#3[#4]{\endgroup
58 \immediate\write-1{Package: #3 #4}%
59 \xdef#1{#4}%
60 }%
61 \else
62 \def\x#1#2[#3]{\endgroup
63 #2[#3]%
64 \ifx#1\relax
65 \xdef#1{#3}%
66 \fi
67 }%
68 \fi
69 \expandafter\x\csname ver@setouterhbox.sty\endcsname
70 \ProvidesPackage{setouterhbox}%
71 [2006/08/26 v1.4 Set hbox in outer horizontal mode (H0)]
\setouterhboxBox The method requires a global box assignment. To be on the safe side, a new box
register is allocated for this global box assignment.
72 \newbox\setouterhboxBox

```

`\setouterhboxFailure` Error message for both plain-TeX and LaTeX

```

73 \def\setouterhboxFailure#1#2{%
74 \begingroup
75 \expandafter\ifx\csname PackageError\endcsname\relax
76 \def\MessageBreak{^^J}%
77 \newlinechar=10\relax
78 \edef\z{#2}%
79 \errhelp\expandafter{\z}%
80 \errmessage{Package setouterhbox Error: #1}%
81 \else
82 \PackageError{setouterhbox}{#1}{#2}%
83 \fi
84 \endgroup
85 }

```

2.2 Main part

eTeX provides much better means for checking error conditions. Thus lines marked by "E" are executed if eTeX is available, otherwise the lines marked by "T" are

used.

```

86 \begingroup\expandafter\expandafter\expandafter\endgroup
87 \expandafter\ifx\csname lastnodetype\endcsname\relax
88   \catcode'T=9 % ignore
89   \catcode'E=14 % comment
90 \else
91   \catcode'T=14 % comment
92   \catcode'E=9 % ignore
93 \fi

```

`\setouterhboxRemove` Remove all kern, glue, and penalty nodes; poor man's version, if ε -TeX is not available

```

94 \def\setouterhboxRemove{%
95 E \ifnum\lastnodetype<11 %
96 E   \else
97 E   \ifnum\lastnodetype>13 %
98 E   \else
99     \unskip\unkern\unpenalty
100 E   \expandafter\expandafter\expandafter\setouterhboxRemove
101 E   \fi
102 E \fi
103 }%

```

`\setouterhbox` Passing the box contents by macro parameter would prevent catcode changes in the box contents like by `\verb`. Also `\bgroup` and `\egroup` does not work, because stuff has to be added at the begin and end of the box, thus the syntax `\setouterhbox{<box number>}\endsetouterhbox` is used. Also we automatically get an environment `setouterhbox` if L^AT_EX is used.

```

104 \def\setouterhbox#1{%
105   \begingroup
106   \def\setouterhboxNum{#1}%
107   \setbox0\vbox\bgroup
108 T   \kern.123pt\relax % marker
109 T   \kern0pt\relax % removed by \setouterhboxRemove
110   \begingroup
111   \everypar{}%
112   \noindent
113 }

```

`\endsetouterhbox` Most of the work is done in the end part, thus the heart of the method follows:

```

114 \def\endsetouterhbox{%
115   \endgroup

```

Omit the first pass to get the penalties of the second pass.

```

116   \pretolerance-1 %

```

We don't want a third pass with `\emergencystretch`.

```

117   \tolerance10000 %
118   \hsize\maxdimen

```

Line is not underfull:

```

119   \parfillskip 0pt plus 1filll\relax
120   \leftskip0pt\relax

```

Suppress underful `\hbox` warnings, is explicit line breaks are used.

```

121   \rightskip0pt plus 1fil\relax
122   \everypar{}%

```

Ensure that there is a paragraph and prevents `\endgraph` from eating terminal glue:

```

123   \kern0pt%
124   \endgraf
125   \setouterhboxRemove

```

```

126 E \ifnum\lastnodetype=1 %
127 E \global\setbox\setouterhboxBox\lastbox
128 E \loop
129 E \setouterhboxRemove
130 E \ifnum\lastnodetype=1 %
131 E \setbox0=\lastbox
132 E \global\setbox\setouterhboxBox=\hbox{%
133 E \unhbox0 %

```

Remove \rightskip, a penalty with -10000 is part of the previous line.

```

134 E \unskip
135 E \unhbox\setouterhboxBox
136 E }%
137 E \repeat
138 E \else
139 E \setouterhboxFailure{%
140 E Something is wrong%
141 E }{%
142 E Could not find expected line.%
143 E \MessageBreak
144 E (\string\lastnodetype: \number\lastnodetype, expected: 1)%
145 E }%
146 E \fi
147 E \setouterhboxRemove
148 T \global\setbox\setouterhboxBox\lastbox
149 T \loop
150 T \setouterhboxRemove
151 T \setbox0=\lastbox
152 T \ifcase\ifvoid0 1\else0\fi
153 T \global\setbox\setouterhboxBox=\hbox{%
154 T \unhbox0 %

```

Remove \rightskip, a penalty with -10000 is part of the previous line.

```

155 T \unskip
156 T \unhbox\setouterhboxBox
157 T }%
158 T \repeat
159 T \ifdim.123pt=\lastkern
160 T \else
161 T \setouterhboxFailure{%
162 T Something is wrong%
163 T }{%
164 T Unexpected stuff was detected before the line.%
165 T }%
166 T \fi
167 T \egroup
168 T \ifcase \ifnum\wd0=0 \else 1\fi
169 T \ifdim\ht0=.123pt \else 1\fi
170 T \ifnum\dp0=0 \else 1\fi
171 T 0 %
172 E \ifnum\lastnodetype=-1 %

```

There was just one line that we have caught.

```

173 \else
174 \setouterhboxFailure{%
175 Something is wrong%
176 }{%
177 After fetching the line there is more unexpected stuff.%
178 E \MessageBreak
179 E (\string\lastnodetype: \number\lastnodetype, expected: -1)%
180 }%
181 \fi
182 E \egroup
183 \expandafter\endgroup

```

```

184 \expandafter\setouterhboxFinish\expandafter{%
185   \number\setouterhboxNum
186 }%
187 }

```

2.3 Environment support

Check `\@currenvir` for the case that `\setouterhbox` was called as environment. Then the box assignment must be put after the `\endgroup` of `\end{...}`.

```

188 \def\setouterhboxCurr{\setouterhbox}
189 \def\setouterhboxLast#1{%
190   \setbox#1\hbox{%
191     \unhbox\setouterhboxBox
192     \unskip % remove \rightskip glue
193     \unskip % remove \parfillskip glue
194     \unpenalty % remove paragraph ending \penalty 10000
195     \unkern % remove explicit kern inserted above
196   }%
197 }

```

`\setouterhboxFinish` #1 is an explicit number.

```

198 \def\setouterhboxFinish#1{%
199   \begingroup\expandafter\expandafter\expandafter\endgroup
200   \expandafter\ifx\csname @currenvir\endcsname\setouterhboxCurr
201     \aftergroup\setouterhboxLast
202     \aftergroup{%
203       \setouterhboxAfter #1\NIL
204     \aftergroup}%
205   \else
206     \setouterhboxLast{#1}%
207   \fi
208 }

```

`\setouterhboxAfter` #1 is an explicit number.

```

209 \def\setouterhboxAfter#1#2\NIL{%
210   \aftergroup#1%
211   \ifx\#2\%
212     \else
213       \setouterhboxReturnAfterFi{%
214         \setouterhboxAfter#2\NIL
215       }%
216   \fi
217 }

```

`\setouterhboxReturnAfterFi` A utility macro to get tail recursion.

```

218 \long\def\setouterhboxReturnAfterFi#1\fi{\fi#1}

```

Restore catcodes we have need to distinguish between the implementation with and without ε -TEX.

```

219 \catcode69=11\relax % E
220 \catcode84=11\relax % T

```

2.4 Option hyperref

```

221 \begingroup
222   \def\x{LaTeX2e}%
223   \expandafter\endgroup
224   \ifx\x\fmtname
225     \else
226       \expandafter\endinput
227   \fi

```

```

\Hy@setouterhbox \Hy@setouterhbox is the internal hook that hyperref uses since 2006/02/12 v6.75a.
228 \DeclareOption{hyperref}{%
229   \long\def\Hy@setouterhbox#1#2{%
230     \setouterhbox{#1}#2\endsetouterhbox
231   }%
232 }

233 \ProcessOptions\relax
234 \</package>

```

3 Installation

CTAN. This package is available on CTAN¹:

[CTAN:macros/latex/contrib/oberdiek/setouterhbox.dtx](#) The source file.

[CTAN:macros/latex/contrib/oberdiek/setouterhbox.pdf](#) Documentation.

Unpacking. The `.dtx` file is a self-extracting `docstrip` archive. The files are extracted by running the `.dtx` through plain- \TeX :

```
tex setouterhbox.dtx
```

TDS. Now the different files must be moved into the different directories in your installation TDS tree (also known as `texmf` tree):

<code>setouterhbox.sty</code>	→	<code>tex/generic/oberdiek/setouterhbox.sty</code>
<code>setouterhbox.pdf</code>	→	<code>doc/latex/oberdiek/setouterhbox.pdf</code>
<code>setouterhbox-example.tex</code>	→	<code>doc/latex/oberdiek/setouterhbox-example.tex</code>
<code>setouterhbox.dtx</code>	→	<code>source/latex/oberdiek/setouterhbox.dtx</code>

If you have a `docstrip.cfg` that configures and enables `docstrip`'s TDS installing feature, then some files can already be in the right place, see the documentation of `docstrip`.

Refresh file databases. If your \TeX distribution (`te \TeX` , `mik \TeX` , ...) rely on file databases, you must refresh these. For example, `te \TeX` users run `texhash` or `mktextlsr`.

3.1 Some details for the interested

Attached source. The PDF documentation on CTAN also includes the `.dtx` source file. It can be extracted by AcrobatReader 6 or higher. Another option is `pdftk`, e.g. unpack the file into the current directory:

```
pdftk setouterhbox.pdf unpack_files output .
```

Unpacking with \LaTeX . The `.dtx` chooses its action depending on the format:

plain- \TeX : Run `docstrip` and extract the files.

\LaTeX : Generate the documentation.

If you insist on using \LaTeX for `docstrip` (really, `docstrip` does not need \LaTeX), then inform the autodetect routine about your intension:

```
latex \install=y\input{setouterhbox.dtx}
```

Do not forget to quote the argument according to the demands of your shell.

¹<http://ftp.ctan.org/tex-archive/>

Generating the documentation. You can use both the `.dtx` or the `.drv` to generate the documentation. The process can be configured by the configuration file `ltxdoc.cfg`. For instance, put this line into this file, if you want to have A4 as paper format:

```
\PassOptionsToClass{a4paper}{article}
```

An example follows how to generate the documentation with pdfL^AT_EX:

```
pdflatex setouterhbox.dtx
makeindex -s gind.ist setouterhbox.idx
pdflatex setouterhbox.dtx
makeindex -s gind.ist setouterhbox.idx
pdflatex setouterhbox.dtx
```

4 References

- [1] Damian Menscher, [news:comp.text.tex](http://groups.google.com/group/comp.text.tex/msg/79648d4cf1f8bc13), *overlong lines in List of Figures*, `<dh058tqbd1@news.ks.uiuc.edu>`, 23rd September 2005. <http://groups.google.com/group/comp.text.tex/msg/79648d4cf1f8bc13>
- [2] David Kastrup, [news:comp.text.tex](http://groups.google.com/group/comp.text.tex/msg/7cf0a345ef932e52), *Re: ANN: outerhbox.sty – collect horizontal material, for unboxing into a paragraph*, `<85y855lrx3.fsf@lola.goethe.zz>`, 7th October 2005. <http://groups.google.com/group/comp.text.tex/msg/7cf0a345ef932e52>
- [3] Michael Downes, *Line breaking in \unhboxed Text*, TUGboat 11 (1990), pp. 605–612.
- [4] Sebastian Rahtz, Heiko Oberdiek: *The hyperref package*; 2006/08/16 v6.75c; CTAN:macros/latex/contrib/hyperref/.

5 History

[2005/10/05 v1.0]

- First version.

[2005/10/07 v1.1]

- Option `hyperref` added.

[2005/10/18 v1.2]

- Support for explicit line breaks added.

[2006/02/12 v1.3]

- DTX format.
- Documentation extended.

[2006/08/26 v1.4]

- Date of `hyperref` updated.

6 Index

Numbers written in *italic* refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; numbers in roman refer to the code lines where the entry is used.

Symbols	M
<code>\</code> 211	<code>\maxdimen</code> 118
	<code>\MessageBreak</code> 76, 143, 178
A	N
<code>\aftergroup</code> 201, 202, 204, 210	<code>\newbox</code> 72
	<code>\newlinechar</code> 77
B	<code>\newsavebox</code> 6
<code>\begin</code> 11, 23	<code>\NIL</code> 203, 209, 214
	<code>\noindent</code> 112
C	<code>\number</code> 144, 179, 185
<code>\catcode</code> 88, 89, 91, 92, 219, 220	
<code>\csname</code> 33, 43, 56, 69, 75, 87, 200	
D	P
<code>\DeclareOption</code> 228	<code>\PackageError</code> 82
<code>\documentclass</code> 2	<code>\PackageInfo</code> 48
<code>\dp</code> 170	<code>\parfillskip</code> 119, 193
	<code>\parindent</code> 8
E	<code>\parskip</code> 9
<code>\empty</code> 37	<code>\penalty</code> 194
<code>\end</code> 26, 29	<code>\pretolerance</code> 116
<code>\endcsname</code> . 33, 43, 56, 69, 75, 87, 200	<code>\ProcessOptions</code> 233
<code>\endgraf</code> 124	<code>\ProvidesPackage</code> 70
<code>\endinput</code> 52, 226	
<code>\endsetouterhbox</code> 114, 230	R
<code>\errhelp</code> 79	<code>\raggedright</code> 12
<code>\errmessage</code> 80	<code>\repeat</code> 137, 158
<code>\everypar</code> 111, 122	<code>\rightskip</code> 121, 192
F	S
<code>\fmtname</code> 224	<code>\sbox</code> 17
	<code>\setbox</code> 107,
H	127, 131, 132, 148, 151, 153, 190
<code>\hbox</code> 132, 153, 190	<code>\setlength</code> 8, 9
<code>\hsize</code> 118	<code>\setouterhbox</code> 104, 230
<code>\ht</code> 169	<code>\setouterhboxAfter</code> 203, 209
<code>\Hy@setouterhbox</code> 228	<code>\setouterhboxBox</code> 72,
I	127, 132, 135, 148, 153, 156, 191
<code>\ifcase</code> 34, 152, 168	<code>\setouterhboxCurr</code> 188, 200
<code>\ifdim</code> 159, 169	<code>\setouterhboxFailure</code> 73, 139, 161, 174
<code>\ifnum</code> .. 95, 97, 126, 130, 168, 170, 172	<code>\setouterhboxFinish</code> 184, 198
<code>\ifvoid</code> 152	<code>\setouterhboxLast</code> 189, 201, 206
<code>\ifx</code> 35,	<code>\setouterhboxNum</code> 106, 185
37, 43, 56, 64, 75, 87, 200, 211, 224	<code>\setouterhboxRemove</code>
<code>\immediate</code> 45, 58 94, 109, 125, 129, 147, 150
	<code>\setouterhboxReturnAfterFi</code> . 213, 218
K	T
<code>\kern</code> 108, 109, 123	<code>\testbox</code> 6, 17, 21, 23, 27
	<code>\tolerance</code> 117
L	U
<code>\lastbox</code> 127, 131, 148, 151	<code>\unhbox</code> . 21, 27, 133, 135, 154, 156, 191
<code>\lastkern</code> 159	<code>\unkern</code> 99, 195
<code>\lastnodetype</code>	<code>\unpenalty</code> 99, 194
... 95, 97, 126, 130, 144, 172, 179	<code>\unskip</code> 99, 134, 155, 192, 193
<code>\leftskip</code> 120	<code>\url</code> 14, 18, 24
<code>\loop</code> 128, 149	<code>\usepackage</code> 3, 4

V		X	
<code>\vbox</code> 107	<code>\x</code> 33, 35, 37, 44, 48, 50, 57, 62, 69, 222, 224
W		Z	
<code>\wd</code> 168		
<code>\write</code> 45, 58	<code>\z</code> 78, 79