How to deal with underfull and overfull boxes

Gady Kozma

29th March 2002

1 What are we talking about?

Like all wordprocessing software, LaTeX, and therefore LyX, will encouter from time to time a situation where it is unable to divide text into lines or divide the paper into pages in an aesthetic manner. Unlike some other wordprocessors, LaTeX will complain about it, by printing messages like overfull hbox and underfull vbox. More notably, its reaction to the first problem is not to create a line with enormous white gaps but to create a line that sticks out of the margin[1].

Unfortunately, LyX doesn't have a very convinient mechanism to find the places in the document where LaTeX thought there were problems of this kind. Thus you must use either of the following methods: choose View ▷ Latex Logfile and search for the error messages, or, usually easier, choose File ▷ Export ▷ Latex and use a terminal to run LaTeX directly, i.e. type the command latex myfile.tex. LaTeX will spew out mysterious error messages, among them the infamous overfull hbox and underfull vbox messages.

Always remember that these problems are **very** sensitive to font, font size, and page size. Therefore if this is not the final version of the document, if at all possible, just ignore them. The time to handle them is when your work has taken on its final form. And don't worry about this too much. It shouldn't take you more than a few minutes work per 10 pages to fix even a complicated document.

Before using all the sophisticated techniques discussed here always remember it might be easier to just change the formulation of the text a little. Hey, don't give me that look! Cosi Fanni Tutti.

2 Overfull hboxes

An "hbox" is an object that LaTeX tries to fill horizontally, typically a line. It's "overfull" when LaTeX cannot push al the things it wants to into the line. The reason is typically a long word or a formula at the end which, if pushed to the next line would make the line too spacey. So LaTeX calls for help. Here is where you step in.

[1] A third difference between LaTeX and some other word processors which is perhaps less notable, but important nonetheless, is the fact that LaTeX breaks lines in a more-than-average intelligent manner. Specifically, if LaTeX encounters a problem in a certain line, it can rearrange all lines leading up to it. As a result, it is unlikely that LaTeX will have problems breaking any line after the second line of a given paragraph.

regular	$(\lbrack\{\vert\}\rbrack)\langle\lfloor\lceil/\uparrow\Uparrow\Downarrow\downarrow\backslash\rceil\rfloor\rangle$
\big	$\big(\big\lbrack\big\{\big\vert\big\}\big\rbrack\big)\big\langle\big\lfloor\big\lceil/\uparrow\Uparrow\Downarrow\downarrow\backslash\big\rceil\big\rfloor\big\rangle$
\Big	$\Big(\Big\lbrack\Big\{\Big\vert\Big\}\Big\rbrack\Big)\Big\langle\Big\lfloor\Big\lceil/\uparrow\Uparrow\Downarrow\downarrow\backslash\Big\rceil\Big\rfloor\Big\rangle$
\bigg	$\bigg(\bigg\lbrack\bigg\{\bigg\vert\bigg\}\bigg\rbrack\bigg)\bigg\langle\bigg\lfloor\bigg\lceil/\uparrow\Uparrow\Downarrow\downarrow\backslash\bigg\rceil\bigg\rfloor\bigg\rangle$
\Bigg	$\Bigg(\Bigg\lbrack\Bigg\{\Bigg\vert\Bigg\}\Bigg\rbrack\Bigg)\Bigg\langle\Bigg\lfloor\Bigg\lceil/\uparrow\Uparrow\Downarrow\downarrow\backslash\Bigg\rceil\Bigg\rfloor\Bigg\rangle$

Table 1: parenthesis and friends in various sizes.

Before you start, check if the complaint is for real. LaTeX will sometimes complain about hardly noticeable problems. If you cannot spot the problem immidiately, one solution is as follows: in the Extra options box under Layout ▷ Document ▷ Document add the word draft. In many document classes (in particular the standard and AMS classes) this will cause a small black box to be printed to the right of the problematic line. If you cannot see the problem — good for you. Ignore it and move on to the next error. Otherwise proceed reading.

If the problem is a long word which LaTeX cannot, or would not, hyphenate, the simplest solution is to add a hyphenation mark where you think the word should be broken. Use Insert ▷ Special Character ▷ Hyphenation Point. If you want the word to be broken but do not want a hyphen to appear (for example, you are breaking a url) you need to change to TeX mode (a.k.a evil red text) — use Layout ▷ Tex Style and write \linebreak[2]. This will force LaTeX to break in that point. Usually, however, you would want to use \linebreak[*n*], which only recommends to LaTeX to break the line there rather than forces it to do so. The number *n* is the level of recommendation, from 0 to 4, with \linebreak[0] being the weakest (and usually what you want) and \linebreak[4] the strongest. Using this method instead of simply \linebreak will mean that you'll have less work the **next** time you change font...

Formulas are more of a problem as LaTeX is typically shy of breaking formulas, and many authors consider breaking formulas across lines an ugly habit[3]. The first option you should consider is probably to make the formula into a display formula (Insert ▷ Display Formula). If you want to keep the formula inline, you can try the \linebreak mechanism above. Notice that to create evil red text in math mode you don't need to select it from a menu, it is activated automatically by the \ sign. In many cases it works inside formulas. The main case where it doesn't is inside grouped parenthesis, by which I mean parenthesis entered using the parenthesis button in the (Edit ▷ Math Panel). LaTeX will not break the line inside a grouped parenthesis no matter what you do. The solution is to replace them with regular parenthesis. You would probably want to enlarge them, though. To do that, just write \big before each parenthesis. If \big is not big enough, just try \Big, \bigg or \Bigg. Table 1 shows how they look.

By the way: I found no gui for entering the symbols $\langle$, $\lfloor$, $\lceil$, \, $\rceil$, $\rfloor$, $\rangle$. I simply enter them using their TeX shortcuts and LyX understands and transforms them to the

[2] Notice that this is **not** the same as Insert ▷ Special Character ▷ Linebreak.

[3] Section 2.1 is for you guys. Naturally, using the techniques there will only aggrevate overfull hbox problems.

appropriate characters automatically. The shortcuts are, respectively, \langle, \lfloor, \lceil, \backslash, \rceil, \rfloor, \rangle.

A final note about this technique: if you use inverted parenthesis, as for example in the old-fashioned style $]0,1[$, LaTeX will do the spacing correctly **only** if you insert them as grouped parenthesis. If you have line-breaking problems with these constructs, and you converted them into non-grouped parenthesis, you will have to tweak the spacing yourself. This can be done by one of the following equivalent methods:

- Hitting C-Space and then, before moving the cursor, hitting Space to shift between various space lengths.
- Choosing from the math panel the small red half-rectangle which leads to a dialogue box allowing you to add a space command
- Using the LaTeX commands \, \: \; and \! for negative space.

Some people claim that LaTeX has a similar problem with spacing around | signs and advocate the use of the AMS commands \lvert and \rvert. (your TeX distribution is likely to have documentation for the amsmath package — try e.g. /usr/share/texmf/doc/latex/amsmath/amsldoc.dvi — where these things are thoroughly explained). Personally I have always had problems actually seeing the difference.

2.1 More on breaking formulas

LaTeX (and TeX) will break inline formulas around equality-like signs, e.g. $=$, $\approx$, $\leq$ etc., and less willingly around binary operators like $+$, $\cdot$, $\wedge$. To make LaTeX less or more likely to break formulas in that way, do the following: in the latex preamble (use Layout ▷ Latex Preamble) add the following commands:

```
\relpenalty=n
\binoppenalty=m
```

where *n* and *m* are two numbers that tell LaTeX how bad is it to break formulas around equality-like signs and binary operators respectively. The default values (for both TeX and LaTeX) are 500 and 700. The value 10000 is used as ∞ so if you set either to 10000 TeX will never break your formulas in that position.

2.2 Display formulas

Display formulas can also be too long, quite easily so in fact, and will create an overfull hbox warning. There isn't much to do about it except transform them into multi-line equations. Multi-line equations are covered quite fully in the user's guide (as of writing this, in section 5.3).

2.3 Underfull hboxes

These are quite rare, but can happen, e.g. when you use forced linebreaks. Here is an example:

> Wow, why is this line like that?
> Oh, because I asked so explicitly. Silly me.

3 Underfull vboxes

A vbox is something that LaTeX tries to fill vertically, typically a page. For some unknown reason, when LaTeX cannot fill the page because the last object is too large, it will not do as it does for lines and go over the margins, but it will prefer to put the offending object on the next page. Hence the error will usually be underfull vbox and the visible symptom will be pages which are too short or that have too much space between paragraphs, formulas etc.

The range of solutions to underfull vboxes is more restricted. Most notably it is to tell LaTeX where to break pages yourself. This is covered in the user's guide under "forcing page breaks" (as of the time of writing, this is subsection 6.3.4). Unfortunately, in many cases the offending object is a multi-line formula, and you cannot enter a pagebreak inside a formula. Thus the simplest solution is to break the formula into two formulas, possibly forcing a pagebreak between them. This should work in most cases.

There is another solution, though, which is rather extreme but might be the right thing for you. It involves using AMS formula environments. I will not explain what they are — this is a rather large topic and in effect they have all kinds of advantages over regular display equations in addition to the ability to break pages inside them. They are covered thoroughly in the AMS documentation I already mentioned above. Unfortunately, LyX has absolutely no support for AMS formula environments, not even in the form of an ugly hack. Therefore proceed as follows:

- Make sure this is the final version of the equation you want.
- Save your file as LaTeX (I'll assume it's called myfile.lyx so the TeX file would be called myfile.tex).
- Open myfile.tex in another editor, find the equation involved, copy it into LyX as text and change it to TeX mode (use Layout ▷ Tex Style).
- Change the equation brackets to whatever you want to use according to the amsmath documetnation which you must read carefully. In particular notice that the align environment needs one & and not two like the eqnarray environment.
- Add a \displaybreak or a \displaybreak[*n*] where you want the page to break. *n* is as for the \linebreak command. Alternatively put the command \allowdisplaybreaks[*n*] in your preamble where the parameter of permissiveness *n* is between 1 (not 0!) and 4.

4 About this document

If you have any comments about this document I would love to hear them. My e-mail is gadykozma@hotmail.com.

This document is in the public domain. Please do with it as you please.