# **DocOnce Description**

### Hans Petter Langtangen<sup>1,2</sup>

<sup>1</sup>Center for Biomedical Computing, Simula Research Laboratory <sup>2</sup>Department of Informatics, University of Oslo

Aug 8, 2014

## 1 What Is DocOnce?

DocOnce is a very simple and minimally tagged markup language that looks like ordinary ASCII text, much like what you would use in an email, but the text can be transformed to numerous other formats, including HTML, Sphinx, LaTeX, PDF, reStructuredText (reST), Markdown, MediaWiki, Google wiki, Creole wiki, blogger.com, wordpress.com, Epytext, and also plain (untagged) text for email. From reST or Markdown you can go to XML, OpenOffice, MS Word, HTML, LaTeX, PDF, DocBook, GNU Texinfo, and more.

DocOnce supports a working strategy of never duplicating information. Text is written in a single place and then transformed to a number of different destinations of diverse type: scientific reports, software manuals, books, thesis, software source code, wikis, blog posts, emails, etc. The slogan is: "Write once, include anywhere".

Here are some DocOnce features:

- DocOnce addresses small and large documents containing text with much computer source code and LTEX mathematics, where the output is desired in different formats such as LTEX, PDFLTEX, Sphinx, HTML, MediaWiki, blogger.com, and wordpress.com. A piece of DocOnce text can enter (e.g.) a classical science book, an ebook, a web document, and a blog post.
- DocOnce targets in particular large book projects where many different pieces of text and software can be assembled and published in different formats for different devices.
- DocOnce enables authors who write for many times of media (blog posts, wikis, LATEX manuscripts, Sphinx, HTML) to use a common source language such that lots of different pieces can easily be brought together later to form a coherent (big) document.

- DocOnce has good support for copying computer code directly from the source code files via regular expressions for the start and end lines.
- DocOnce first runs two preprocessors (Preprocess and Mako), which allow programming constructs (includes, if-tests, function calls) as part of the text. This feature makes it easy to write one text with different flavors: long vs short text, Python vs Matlab code examples, experimental vs mature content.
- DocOnce can be converted to plain *untagged* text, often desirable for email and computer code documentation.
- DocOnce markup does include tags, so the format is more tagged than Markdown, but less than reST, and very much less than LATEX and HTML.
- Compared to the related tools Sphinx and Markdown, DocOnce allows more types of equations (especially systems of equations with references), has more flexible inclusion of source code, integrates preprocessors, has special support for exercises, and produces cleaner Lagrange ATEX and HTML output.

History. The DocOnce development started in 2006 at a time when most popular markup languages used quite some tagging (LATEX, reStructuredText, HTML). Later, almost untagged markup languages, especially Markdown and its sisters MultiMarkdown and Pandoc-extended Markdown, became popular. DocOnce looks much like Markdown and is in particular close to the functionality and nature of MultiMarkdown. The advantage of DocOnce, however, is a series of features for supporting both small and large documents (books in particular) with much mathematics and computer code. DocOnce can also output Sphinx (not supported by Pandoc or MultiMarkdown), a format that is very attractive for presenting scientific material on the web. A recent dialect of DocOnce allows Markdown syntax, extended with DocOnce syntax as you like, as valid input to the DocOnce translator.

**Disclaimer.** DocOnce applies *text transformations*, mostly via regular expressions. This is not a fool-proof method of translation compared to real parsing. The possibility for tweaking the layout in the DocOnce document is obviously limited (compared to LATEX and HTML in particular) since the text can go to all sorts of sophisticated markup languages. This can be compensated, however, by clever use of the programmable preprocessors and by automatic editing of the output (via regular expressions).

#### Contents

1 What Is DocOnce?

	1.1	Demos	4
2	2.1	DocOnce Markup Language Lists	<b>5</b>
	2.2	Special lines	6
	2.3	Heading with title and author(s)	6
	2.4	Table of contents	7
	2.5	Section headings	7
	2.6	Figures	8
	2.7	Movies	10
	2.8	Copying Computer Code from Source Files	13
	2.9		13
	2.10		14
			21
			22
			22
			26
			27
			29
	2 17		31
			34
			38
			39
			39
			43
			13
			44
			16
			+0 46
			+0 46
			+0 18
	2.20	Emacs Doconce Formatter	łC
3	Writ	ing Slides	18
	3.1		19
	3.2	HTML5 Slides	_
	3.3		50
	0.0	Elex beamer olides	,
4	Mak	o Programming	50
	4.1	The Basics of Mako	50
	4.2	Debugging Python code in Mako	51
	4.3		52
	4.4		53
	4.5		58
	4.6		30
	47		รด

5	From DocOnce to Other Formats	6	0
	5.1 Generating a makefile	. 6	31
	5.2 Preprocessing	. 6	31
	5.3 Removal of inline comments	_	2
	5.4 Notes	. 6	2
	5.5 Demo of different formats		2
	5.6 HTML	. 6	2
	5.7 Blog Posts	_	34
	5.8 Pandoc and Markdown	. 6	6
	5.9 LATEX	_	8
	5.10 PDFLaTeX		2
	5.11 XeLaTeX		'3
	5.12 From PDF to e-book formats		'3
	5.13 Plain ASCII Text		'3
	5.14 reStructuredText	. 7	'3
	5.15 Sphinx		4
	5.16 Wiki Formats		'8
	5.17 Google Docs		'9
	5.18 Tweaking the DocOnce Output	. 7	'9
6	Options for the doconce commands	8	0
	6.1 doconce format command-line options	. 8	0
7	Installation of DocOnce and its Dependencies	9	0
	7.1 DocOnce	. 9	0
	7.2 Dependencies		1
	7.3 Quick Debian/Ubuntu Install		6
8	Basic Parsing Ideas	9	16
•	8.1 Typesetting of Function Arguments, Return Values, and Variable	•	7
9	References	0	8
J		9	0

### 1.1 Demos

The current text is generated from a DocOnce format stored in the directory doc/manual/manual.do.txt

file in the DocOnce source code tree. Here you can run a make.sh script to generate a lot of different formats: HTML, LaTeX, plain text, etc., stored in the subdirectory demo. To successfully run this script and compile all the different versions, you must have quite a lot of software installed, see the Section 7 for detailed information. However, the most commonly used formats of this manual are fortunately also accessible through the web: HTML, Sphinx, and PDF.

DocOnce supports a lot of different layouts of documents, see the collection of various versions of a scientific report.

Another demo is found in

doc/tutorial/tutorial.do.txt

In the tutorial directory there is also a make.sh file producing a lot of formats in the subdirectory demo. The three most common formats can be found on the web: HTML, Sphinx, and PDF.

# 2 The DocOnce Markup Language

The DocOnce format introduces four constructs to markup text: lists, special lines, inline tags, and environments.

#### 2.1 Lists

An unordered bullet list makes use of the \* as bullet sign and is indented as follows

- \* item 1
- \* item 2
  - \* subitem 1, if there are more lines, each line must be intended as shown here
  - \* subitem 2, also spans two lines
- \* item 3

This list gets typeset as

- item 1
- item 2
  - subitem 1, if there are more lines, each line must be intended as shown here
  - subitem 2, also spans two lines
- item 3

In an ordered list, each item starts with an o (as the first letter in *ordered*):

```
o item 1
o item 2
 * subitem 1
 * subitem 2
o item 3
```

resulting in

- 1. item 1
- 2. item 2
  - subitem 1
  - subitem 2
- 3. item 3

Ordered lists cannot have an ordered sublist, i.e., the ordering applies to the outer list only.

In a description list, each item is recognized by a dash followed by a keyword followed by a colon:

```
- keyword1: explanation of keyword1
```

```
- keyword2: explanation
  of keyword2 (remember to indent properly
  if there are multiple
  lines)
```

The result becomes

keyword1: explanation of keyword1

**keyword2:** explanation of keyword2 (remember to indent properly if there are multiple lines)

## 2.2 Special lines

The DocOnce markup language has a concept called *special lines*. Such lines starts with a markup at the very beginning of the line and are used to mark document title, authors, date, sections, subsections, paragraphs, figures, movies, etc.

## 2.3 Heading with title and author(s)

Lines starting with TITLE:, AUTHOR:, and DATE: are optional and used to identify a title of the document, the authors, and the date. The title is treated as the rest of the line, so is the date, but the author text consists of the name and associated institution(s) with the syntax

```
name at institution1 and institution2 and institution3
```

The at with surrounding spaces is essential for adding information about institution(s) to the author name, and the and with surrounding spaces is essential as delimiter between different institutions. An email address can optionally be included, using the syntax

name Email: somename@site.net at institution1 and institution2

Multiple authors require multiple AUTHOR: lines. All information associated with TITLE: and AUTHOR: keywords must appear on a single line. Here is an example:

TITLE: On an Ultimate Markup Language
AUTHOR: H. P. Langtangen at Center for Biomedical Computing, Simula Research Laboratory & Dept. of Informat
AUTHOR: Kaare Dump Email: dump@cyb.space.com at Segfault, Cyberspace Inc.
AUTHOR: A. Dummy Author
DATE: November 9, 2016

Note how one can specify a single institution, multiple institutions (with & as separator between institutions), and no institution. In some formats (including rst and sphinx) only the author names appear. Some formats have "intelligence" in listing authors and institutions, e.g., the plain text format:

```
Hans Petter Langtangen [1, 2]
Kaare Dump (dump@cyb.space.com) [3]
A. Dummy Author
```

- [1] Center for Biomedical Computing, Simula Research Laboratory
- [2] Department of Informatics, University of Oslo
- [3] Segfault, Cyberspace Inc.

Similar typesetting is done for LaTeX and HTML formats. The current date can be specified as today.

#### 2.4 Table of contents

A table of contents can be generated by the line

TOC: on

This line is usually placed after the DATE: line. The value off turns off the table of contents.

## 2.5 Section headings

Section headings are recognized by being surrounded by equal signs (=) or underscores before and after the text of the headline. Different section levels are recognized by the associated number of underscores or equal signs (=):

- 9 = characters for chapters
- 7 for sections
- 5 for subsections
- 3 for subsubsections
- 2 *underscrores* (only! it looks best) for paragraphs (paragraph heading will be inlined)

Headings can be surrounded by as many blanks as desired.

DocOnce also supports abstracts. This is typeset as a paragraph, but *must* be followed by a section heading (everything up to the first section heading is taken as part of the text of the abstract).

Here are some examples:

```
__Abstract.__ The following text just attempts to exemplify various section headings.

Appendix is supported too: just let the heading start with ''Appendix: '' (this affects only 'latex' output, where the appendix formatting is used - all other formats just leave the heading as it is written).

====== Example on a Section Heading ======

The running text goes here.

=== Example on a Subsection Heading ===

The running text goes here.

=== Example on a Subsubsection Heading ===

The running text goes here.

__A Paragraph.__ The running text goes here.
```

## 2.6 Figures

Figures are recognized by the special line syntax

```
FIGURE: [filename, height=400 width=600 frac=0.8] caption
```

The filename can be without extension, and DocOnce will search for an appropriate file with the right extension. If the extension is wrong, say .eps when requesting an HTML format, DocOnce tries to find another file, and if not, the given file is converted to a proper format (using ImageMagick's convert utility).

The height, width, and frac keywords can be included if desired and may have effect for some formats: the height and width are used for output in the formats html, rst, sphinx, while the frac specification is used for latex and pdflatex to specify the width of the image as a fraction of the text width.

The figure caption is optional. If omitted, the figure appears "inline" in the text without any figure environment in LaTeX formats or HTML. The caption may contain a label for referencing the figure.

#### Warning.

Note the comma between the filename and the figure size specifications and that there should be no space around the = sign. This syntax must be strictly followed.

Note also that, like for TITLE: and AUTHOR: lines, all information related to a figure line *must be written on the same line*. Introducing newlines in a long caption will destroy the formatting (only the part of the caption appearing on the same line as FIGURE: will be included in the formatted caption).

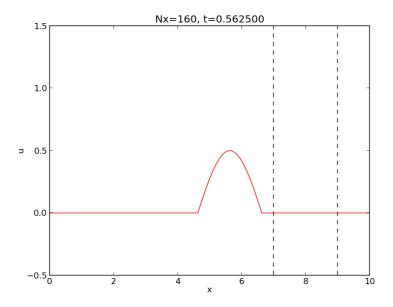


Figure 1: A wave.

Combining several image files into one, in a table fashion, can be done by the montage program from the ImageMagick suite:

```
montage -background white -geometry 100% -tile 2x \ file1.png file2.png ... file4.png result.png
```

The option  $-tile \ XxY \ gives \ X \ figures in the horizontal direction and Y in the vertical direction (tile <math>2x$  means two figures per row and  $-tile \ x2$  means two rows).

The montage program is only appropriate for bitmap images (PNG, JPEG, GIF, TIFF). Images in the PDF format should be mounted together using pdftk (to combine images to one file), pdfnup (to align them in tabular format), and pdfcrop (to remove surrounding whitespace):

```
Terminal> pdftk file1.pdf file2.pdf ... file4.pdf output tmp.pdf
Terminal> pdfnup --nup 2x2 tmp.pdf # output in tmp-nup.pdf
Terminal> pdfcrop tmp-nup.pdf result.png # output in FE1.png
```

#### 2.7 Movies

Movies/videos are inserted using the MOVIE: keyword. This feature works well for the latex, html, rst, and sphinx formats. Other formats try to generate some HTML file and link to that file for showing the movie.

**The Basic Command.** As with FIGURE, the MOVIE command expands just one line and is of the form

MOVIE: [filename, height=xxx width=yyy] possible caption

Note that there must be a blank line after every MOVIE: command. The width and height parameters are not required, but leaving them out may lead to movie sizes you do not want.

Here is a movie in the MPEG format:

mov/wave.mpeg

realized by the specific command

MOVIE: [mov/wave.mpeg, width=600 height=470]

And here is a movie in the Ogg format:

mov/wave.ogg

Movie 2: A movie in Ogg format.

A URL works too as movie address:

MOVIE: [http://hplgit.github.io/animate/doc/pub/mov-animate/demo.ogg] Ogg movie in cyberspace.

http://hplgit.github.io/animate/doc/pub/mov-animate/demo.ogg

Movie 3: Ogg movie in cyberspace.

#### Important.

Movies will not work properly in sphinx format unless they are located in a directory (tree) with a name starting with mov. Make it a habit to place figures in fig-X and movies in mov-X directories, where X is a short logical name for the current document (or let the names of the directories be just fig and mov).

MP4, WebM, and Ogg Movies in HTML. If a movie is in Ogg, MP4, or WebM format, and the output format is HTML, DocOnce will check if the movie file is also in the other formats among Ogg, MP4, and WebM, and include these as well such that the movie has backup formats in case the browser does not support a particular format. Providing a movie in Ogg, MP4, and WebM format is therefore the safest way to ensure that the movie can be played in any browser on any device.

#### Notice.

If you specify a movie in Ogg or WebM format and it also exists in MP4 format, the MP4 format will be loaded first. To avoid having alternative movie formats in HTML, use the --no\_mp4\_webm\_ogg\_alternatives command-line option when running doconce format.

Movie Handling in Various Formats. Movies are easiest shown in the HTML format. The reST and Sphinx formats apply the same raw HTML code as the HTML format and therefore have the same capabilities. The Latex\_movie=.... Proper values are listed below.

- 1. href: the \href{run:file}{link} is used for all movies (default).
- media9: the media9 package is used for Flash and MP4 movies, movie15 for MPEG and AVI files, and a simple \href{run:file}{link} command for other formats. Only Acrobat Reader supports displaying these type of movies.
- 3. multimedia: the \movie command (known from LATEX Beamer) is used for movies.
- 4. movie15: the old movie15 package is used.

For all other formats, an HTML file that acts as a movie player is generated and linked from the output document. This movie player has essentially the same code as the HTML format would have, except that the video tag is not used, only the embed tag. Some wiki types do have support for videos, e.g., Wikipedia can work with Ogg files, but DocOnce has not yet implemented robust schemes for anything but LaTeX, HTML, and Sphinx output.

**YouTube and Vimeo Movies.** Many publish their scientific movies on YouTube or Vimeo, and DocOnce recognizes YouTube and Vimeo URLs as movies. When the output from DocOnce is an HTML file, the movie will be embedded, otherwise a URL to the YouTube or Vimeo page is inserted. You should equip the MOVIE: command with the right width and height of *embedded* YouTube and Vimeo movies. The recipe goes as follows:

- 1. click on Share (and on YouTube then Embed)
- 2. note the height and width of the embedded movie

A typical MOVIE command with a YouTube movie is then

MOVIE: [http://www.youtube.com/watch?v=sI2uCHH3qIM, width=420 height=315]

MOVIE: [http://vimeo.com/55562330, width=500 height=278] CFD.

**Animation Based on Filename Generators.** It is possible to define a movie from a set of files, usually plot files, which can be shown in sequence to create an animation. If the files are local on the computer, one can specify them by a simple Unix wildcard notation, as in

MOVIE: [../experiments/frame\_\*.png]

Output in the HTML, reST, and Sphinx formats will make use of inline JavaScript code to show the frames in sequence. LaTEX employs the animate package for the same purpose. Other formats generates a file (movie\_playerX.py, where X is a number) containing the HTML code with JavaScript to show and control the animation. The DocOnce document has a link to this movie viewer.

There is an alternative syntax to the Unix wildcard notation:

MOVIE: [../experiments/frame\_%04d.png:0->320]

The filename is specified via printf syntax (typically the same syntax as used to generate the individual frame files). The postfix :0->320 specifies the lower and upper limit of the counter that is used in the printf specification %04d. This latter syntax must be used if the plot files reside on some web server, e.g.,

MOVIE: [http://some.where.net/experiments/frame\_%04d.png:0->320]

Here is an example:

http://hplgit.github.io/animate/doc/pub/mov-animate/frames/frame\_%04d.png:0->320: load movie\_player1.html into a browser

#### Recommendations.

It is challenging to write robust DocOnce code with movies. The recommended formats in HTML are MP4, WebM, and Ogg. One should preferably make all three. These also works in reST and Sphinx.

The filename generation works very well in LaTeX, while true movie formats pose big challenges. On Linux systems, the media9 does not work well because a proper Flash player for embedding in the PDF file is not always available. The movie15 package also leads to problems because Acrobat Reader depends on an external player to show the files, and the correct plugins to launch players with support for a given format

are not trivial to install. Even the plain href{run:file} command relies on an external player and not all formats will be supported on a given computer.

To have really robust code, use filename generators and not movie files.

```
MOVIE: [../experiments/frame_*.png]
```

One can, of course, write flexible DocOnce code and decide at run time if HTML output should have movie files or filename generators. A relevant snippet using Mako is

```
% FORMAT in ("latex", "pdflatex") or HTMLMOVIE == "files":
MOVIE: [../experiments/frame_*.png]
% else:
MOVIE: [../experiments/movie.ogg]
% endif
```

With the -DHTMLMOVIE=files flag, animation of individual files will be performed, while any other value than files leads to use of the movie.ogg in all but LATEX formats. In HTML one will try to load movie.mp4 (if it exists) and then movie.webm (if it exists) and then finally movie.ogg.

There is no way to control the number of frames per second in LaTeX animations based on filename generators such as myframes\*.png. However, with a little auto editing in a script one can control the frame rates of the various movies. The rate is specified as 2 in lines on the form

```
\begin{animateinline}[controls,loop]{2} % frames: f000.png -> f098.png
```

Setting the rate to 12 instead for this particular movie based on the f%03.png files, the following doconce subst command does the job in a script:

```
doconce subst ',loop]{2}( .+: f000)' ',loop{12}\g<1>' mydoc.do.txt
```

### 2.8 Copying Computer Code from Source Files

Another type of special lines starts with @@@CODE and enables copying of computer code from a file directly into a verbatim environment, see Section 2.18 below.

### 2.9 Inserting the Output from Operating System Commands

When DocOnce is used to document computer program and results from computer code it is important to ensure that the document contains the latest version of the code and the corresponding output. The former is handled by the

@@@CODE directive, while the latter has its own directive @@@OSCMD. The syntax reads

```
@@@OSCMD cmd
```

where cmd is any text that can be run in the operating system. The output is copied into the DocOnce source. For example,

```
@@@OSCMD python -c 'print "Hello,\nWorld!"'
results in
Terminal> python -c 'print "Hello,\nWorld!"'
Hello,
World!
```

There is a command-line option --os\_prompt= that can be used to set the (terminal) prompt that prefixes the command:

- --os\_prompt=None results in no prompt, just the command.
- --os\_prompt=nocmd results in no prompt and no command, just the output.
- --os\_prompt=Terminal> is the default setting (as in the example above).

## 2.10 Inline Tagging

DocOnce supports tags for *emphasized phrases*, **boldface phrases**, and <code>verbatim text</code> (also called type writer text, for inline code), colored words, plus LaTeX/TeX inline mathematics, such as  $\nu = \sin(x)$ . Links are easy to define, either with a text or just a plain <code>http://google.com</code>. Also a non-breaking space (to avoid linebreak), linebreak,

m-dash (as in m—dash), and horizontal rule can be specified (below).

#### Limitation of inline tagging.

Since DocOnce applies regular expressions to recognize inline tagging, there might be cases where the tags are not correctly interpreted and translated. Fortunately, such pitfalls are easily circumvented. The troubleshooting document shows some examples.

**Emphasized Words.** Emphasized text is typeset inside a pair of asterisk, and there should be no spaces between an asterisk and the emphasized text, as in

```
*emphasized words*
```

Boldface font is recognized by an underscore instead of an asterisk:

```
_several words in boldface_ followed by *ephasized text*.
```

The line above gets typeset as **several words in boldface** followed by *ephasized text*. One should only have pure text (no mathematical formulas) between the boldface or emphasize markers, and no leading or trailing blanks (with such blanks, the text will not be recognized as boldface or emphasize).

Colored text is formatted as

```
some text color{red}{more text in red}
```

**Inline Verbatim Text.** Verbatim text, typically used for short inline code, is typeset between backticks:

```
'call myroutine(a, b)' looks like a Fortran call while 'void myfunc(double *a, double *b)' must be C.
```

The typesetting result looks like this: call myroutine(a, b) looks like a Fortran call while void myfunc(double \*a, double \*b) must be C. Note that there must be no leading or trailing spaces inside the backticks.

It is recommended to have inline verbatim text on the same line in the DocOnce file, because some formats (LATEX in combination with the ptex2tex program (but not doconce pretex)) will have problems with inline verbatim text that is split over two lines.

#### Notice.

Watch out for mixing backticks and asterisk (i.e., verbatim and emphasized code): the DocOnce interpreter is not very smart in detecting such errors. A missing backtick will also quickly create strange output. If you suspect inline code to be the source of problems in the final format, examine the DocOnce source and the output.

**Links to Web Addresses.** Web addresses with links are typeset as

```
some URL like "Search Google": "http://google.com".
```

which appears as some URL like Search Google. The space after colon is optional, but it is important to enclose the link and the URL in double quotes.

To have the URL address itself as link text, put an "URL" or URL before the address enclosed in double guotes:

```
Click on this link: URL: "https://github.com/hplgit/doconce".
```

which gets rendered as Click on this link: https://github.com/hplgit/doconce.

(There is also support for lazy writing of URLs: any http or https web address with a leading space and a trailing space, comma, semi-colon, or question mark (but not period!) becomes a link with the web address as link text.)

**Mail Addresses.** Links that launches a mail to a specified address is written as ordinary URLs, typically as

```
Send "mail": "mailto:hpl@simula.no"
# Alternative:
to "'hpl@simula.no'": "mailto:hpl@simula.no".
```

which appears as Send mail to hpl@simula.no.

**Links to Local Files.** Links to files ending in .txt, .html, .pdf, .py, .f, .f77, .f90, .f95, .sh, .csh, .ksh, .zsh, .c, .cpp, .cxx, .pl, and .java follows the same setup:

```
see the "DocOnce Manual": "manual.do.txt".
```

which appears as see the DocOnce Manual. However, linking to local files like this needs caution:

- In the html format the links work well if the files are supplied with the .html with the same relative location.
- In the latex and pdflatex formats, such links in PDF files will unless the
  .tex file has a full URL specified through a \hyperbaseurl command
  and the linked files are located correctly relative to this URL. Otherwise
  full URL must be used in links.
- In the sphinx format, links to local files do not work unless the files reside in a static directory (a warning is issued about this).

As a consequence, we strongly recommend that one copies the relevant files to a \_static or \_static-name directory and makes links to files in this directory only (name is the nickname of the DocOnce document, usually the name of the parent directory or main document). Other links to files should use the full URL. If DocOnce is used for HTML output only, then plain links to local files work fine.

If you want a link to a local source code file and have it viewed in the browser rather than being downloaded, we recommend to transform the source code file to HTML format by running pygmentize, e.g.,

```
Terminal> pygmentize -l bash -f html -0 full,style=emacs \
-o _static/make.sh.html subdir/make.sh
```

Then you can link to \_static/make.sh.html instead of subdir/make.sh. Here is an example where the reader has the file available as src/myprog.py in her software and the document links to \_static/myprog.py:

```
See the code URL: "src/myprog.py" ("view: "_static/myprog.py.html").
```

Links to files with other extensions are typeset with *the filename as link text*. The syntax consists of the keyword URL, followed by a colon, and then the filename enclosed in double guotes:

```
URL: "manual.html"
```

resulting in the link manual.html.

**Quotes.** Quotations employ either the emphasized font or double quotation marks. In the latter case, one should not use the character " but rather the (LATEX-inspired) construction with double backticks and two single quotes:

```
This is a sentence with "words to be quoted".
```

To candidates for find double quotes that should be transferred to the above type of quotation (which is a common mistake), one can run a regular expression search like

```
Terminal> find . -name '*.do.txt' -exec grep -E \
   '[^("]"[A-Za-z0-9 ,]+" *[^:')"]' {} \; -print
```

This search may give many false hits as double quotes are frequently used in computer code and preprocessor instructions (URLs and hyperlinks should not give hits in the above regular expressions).

**Non-Breaking Space.** The non-breaking space character is tilde:

```
Here comes a long line with a specification of a number with unit at the end, which is an example that requires a "non-breaking space character": "http://en.wikipedia.org/wiki/Non-breaking_space": 7.4~km is traveled in~$7.4/5.\alpha0 prox 1.345$~s. (Computer code, where the tilde has a meaning, as in 'y = \alpha1, is not affected. The non-breaking character only works between characters, numbers, and the math dollar sign.)
```

This is rendered as

Here comes a long line with a specification of a number with unit at the end, which is an example that requires a non-breaking space character: 7.4 km is traveled in  $7.4/5.5 \approx 1.345$  s. (Computer code, where the tilde has a meaning, as in y = ~x, is not affected. The non-breaking character only works between characters, numbers, and the math dollar sign.)

Horizontal rule. A horizontal rule for separating content vertically, like this:

is typeset as four or more hyphens on a single line:

-----

**em-dash.** The latex, pdflatex, sphinx, and html formats support em-dash, indicated by three hyphens: ---. The em-dash has two applications: 1) as alternative to a hyphen with space around in a sentence—except that there are no spaces around the em-dash, and 2) origin of quotes, where there is no space between the end of the quote and the m-dash:

Premature optimization is the root of all evil.— Donald Knuth.

The associated DocOnce source reads

```
!bquote
*Premature optimization is the root of all evil.*--- Donald Knuth.
!equote
```

Ampersand. An ampersand, as in Guns & Roses or Texas A&M, is written as a plain & with space(s) on both sides. Single upper case letters on each side of &, as in Texas A & M, remove the spaces and result in Texas A&M, while words on both sides of &, as in Guns & Roses, preserve the spaces: Guns & Roses. Failing to have spaces before and after & will result in wrong typesetting of the ampersand in the html, latex, and pdflatex formats.

**Footnotes.** Typesetting of footnotes employs a common email or Extended Markdown syntax:

Footnotes are typeset according to the output format [^typesetting]. The syntax is optional spaces, opening bracket, hat, a name without spaces [^remedy-for-name-with-spaces], and closing bracket. The logical name of the footnote is not used in LaTeX, HTML, reStructuredText, or Sphinx, because these languages employ numbered footnotes. Other formats employ the logical name.

[^typesetting]: Typesetting of the footnote depends on the format. Plain text does nothing, LaTeX removes the definition and inserts the footnote as part of the LaTeX text. reStructuredText and Sphinx employ a similar type of typesetting as Extended Markdown and DocOnce, and in HTML we keep the same syntax, just displayed properly in HTML.

Footnotes are preferably defined after the paragraph they are used. The definition is the footnote syntax (some optional space, bracket, hat, name, bracket) followed by colon and a text.

[ $^{\text{remedy-for-name-with-spaces}}$ ]: Just put in dashes or underscores in case of spaces.

The text above looks as follows.

Footnotes are typeset according to the output format<sup>1</sup>. The syntax is optional spaces, opening bracket, hat, a name without spaces<sup>2</sup>, and closing bracket. The logical name of the footnote is not used in LATEX, HTML, reStructuredText, or Sphinx, because these languages employ numbered footnotes. Other formats employ the logical name.

Footnotes are preferably defined after the paragraph they are used. The definition is the footnote syntax (some optional space, bracket, hat, name, bracket) followed by colon and a text.

**Inline Comments.** DocOnce also supports inline comments in the text:

```
[name: comment]
```

where name is (e.g.) the name of the author of the comment, and comment is a plain text text. Note that *there must be a space after the colon*, otherwise the comment is not recognized. The name can contain upper and lower case characters, digits, single quote, + and -, as well as space. Next is an example. **hpl's comment 1**: Inline comments can span several lines, if desired.

The name and comment are visible in the output unless doconce format is run with a command-line argument --skip\_inline\_comments (see Section 5 for an example). Inline comments are helpful during development of a document since different authors and readers can comment on formulations, missing points, etc. All such comments can easily be removed from the .do.txt file by doconce remove\_inline\_comments (see Section 5).

Inline comments are typeset in a simple way: boldface name, a numbering of the comment, and then the comment, all in red and in parenthesis. However, with the --latex\_todonotes option, LaTEX will apply the todonotespackage to typeset the comments in very visible color boxes.

**Inline Comments for Editing.** Inline comments can also be used to markup editing of the text. The following syntax is supported:

```
[add: ,]
[add: .]
[add: ;]
[del: ,]
[del: ,]
[del: .]
[del: .]
[del: some text]
[del: some text]
[edit: some text -> some replacement for text]
[name: some text -> some replacement for text]
```

¹Typesetting of the footnote depends on the format. Plain text does nothing, LaTeX removes the definition and inserts the footnote as part of the LaTeX text. reStructuredText and Sphinx employ a similar type of typesetting as Extended Markdown and DocOnce, and in HTML we keep the same syntax, just displayed properly in HTML.

<sup>&</sup>lt;sup>2</sup>Just put in dashes or underscores in case of spaces.

That is, one can add, delete, and replace text, and adding or deleting a comma, period, or semicolon leads to special typesetting where such a small edit is highlighted. Below is an example of a text with inline editing.

Originally, we have the text

```
First consider a quantity Q. Without loss of generality, we assume Q>0. There are three, fundamental, basic property of Q.
```

Then, some reader wants to change this text and explicitly demonstrate what is deleted, added, and replaced (as when using track changes in Microsoft Word). The use of the add, del, and replacement construction with -> may look as follows.

```
First[add: ,] consider [edit: a quantity -> the flux]
[del: $Q$. Without loss of generality,
we assume] $Q>0$. There are three[del: ,] fundamental[del: , basic]
[edit: property -> properties] of $Q$. [add: These are not
important for the following discussion.]
```

The text gets rendered as

First, (edit 2: add comma) consider (edit 3:) a quantity the flux (edit 4:) Q. Without loss of generality, we assume Q>0. There are three (edit 5: delete comma) fundamental(edit 6:) , basic (edit 7:) property properties of Q. (edit 8:) These are not important for the following discussion.

Such inline comments with edits are only given special typesetting in the output formats latex, pdflatex, html, and sphinx. Otherwise, just the DocOnce syntax is shown (but that is also quite readable as edit instructions.)

The editing implied by the edit comments can be implemented in the DocOnce file by the command

```
Terminal> doconce apply_edit_comments mydoc.do.txt
```

**Forced Line Breaks.** By ending a line with linebreak> the output format has a forced linebreak at this point. This can be used to typeset poems, songs (if not in a verbatim block), or the origin of quotes. Here is an example:

```
!bquote

*Program writing is substantially more demanding than book

writing. Why is it so? I think the main reason is that a larger

attention span is needed when working on a large computer program

than when doing other intellectual tasks.* linebreak>

Donald Knuth cite[p. 18]{Knuth85}, computer scientist, 1938-.

!equote
```

#### is rendered as

Program writing is substantially more demanding than book writing. Why is it so? I think the main reason is that a larger attention span is needed when working on a large computer program than when doing other intellectual tasks.

Donald Knuth [1, p. 18], computer scientist, 1938-.

Inline Mathematics. Inline mathematics is written as in LaTeX, i.e., inside dollar signs. Many formats leave this syntax as it is (including to dollar signs), hence nice math formatting is only obtained in LaTeX, HTML, MediaWiki, and Sphinx (Epytext has some inline math support that is utilized). However, mathematical expressions in LaTeX syntax often contains special formatting commands, which may appear annoying in plain text. DocOnce therefore supports an extended inline math syntax where the writer can provide an alternative syntax suited for formats close to plain ASCII:

```
Here is an example on a linear system {\footnote{$\{\footnote{$\{\footnote{$\{\}\}}}}= \{\footnote{$\{\}\}}\}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}\}}} = {\footnote{$\{\}\}}} = {\
```

That is, we provide two alternative expressions, both enclosed in dollar signs and separated by a pipe symbol, the expression to the left is used in formats with LaTeX support (latex, pdflatex, html, sphinx, mwiki), while the expression to the right is used for all other formats. The above text is typeset as "Here is an example on a linear system  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , where  $\mathbf{A}$  is an  $n \times n$  matrix, and  $\mathbf{x}$  and  $\mathbf{b}$  are vectors of length n."

#### 2.11 Comments

Comments intended to be (sometimes) visible in the output document and read by readers are known as *inline comments* in DocOnce and described in Section 2.10.

Here we address comments in the DocOnce source file that are not intended to be visible in the output document. Basic comment lines start with the hash #:

```
#
# Here are some comment lines that do not affect any formatting.
# These lines are converted to comments in the output format.
#
```

Such comment lines may have some side effects in the rst and sphinx formats because following lines are taken as part of the comment if there is not a blank line after the comment.

The Mako preprocessor supports comments that are filtered out *before* DocOnce starts translating the document. Such comments are very valuable as they will never interfere with the output format and they are only present in the DocOnce source. Mako has two types of comments: lines starting with a double hash ## and lines enclosed by the <%doc> (beginning) and <%doc/> (closing) tags.

If you need a lot of comments in the DocOnce file, consider using Mako comments instead of the single hash, unless you want the comments to be in the source code of the output document.

To comment out or remove large sections, consider using the Preprocess preprocessor and an if-else block with a variable that is undefined (typically something like a test # #ifdef EXTRA in Preprocess).

## 2.12 Cross-Referencing

References and labels are supported. The syntax is simple:

```
label{section:verbatim}  # defines a label
For more information we refer to Section ref{section:verbatim}.
```

This syntax is close that that of labels and cross-references in LaTeX. When the label is placed after a section or subsection heading, the plain text, Epytext, and StructuredText formats will simply replace the reference by the title of the (sub)section. All labels will become invisible, except those in math environments. In the rst and sphinx formats, the end effect is the same, but the "label" and "ref" commands are first translated to the proper reST commands by doconce format. In the HTML and (Google Code) wiki formats, labels become anchors and references become links, and with LaTeX "label" and "ref" are just equipped with backslashes so these commands work as usual in LaTeX.

It is, in general, recommended to use labels and references for (sub)sections, equations, and figures only. By the way, here is an example on referencing Figure 1 (the label appears in the figure caption in the source code of this document). Additional references to Sections 2.19 and 2.20 are nice to demonstrate, as well as a reference to equations, say (??)-(??). A comparison of the output and the source of this document illustrates how labels and references are handled by the format in question.

Hyperlinks to files or web addresses are handled as explained in Section 2.10.

## 2.13 Generalized Cross-Referencing

Sometimes a series of individual documents may be assembled to one large document, typically a book. In the book one wants to make cross references between chapters and sections, while these become references to external documents when the chapters (or sections) are compiled as stand-alone documents. For example, one can in a DocOnce file file1.do.txt have text like

```
...as shown in Section ref{sec:eqs}.
```

with \label{sec:eqs} defined in another file file2.do.txt. If file1.do.txt and file2.do.txt are combined to a single document, the reference is treated correctly, but if file1.do.txt is compiled as a single document, the label sec:eqs becomes undefined. Then one would instead write

```
...as shown in the document "Mathematical Equations": "http://some.net/doc/matheqs.html" cite{math_eqs_2020}.
```

However,  $\LaTeX$  has functionality for referring to labels in external documents: these work fine when the xr package is used and the external documents are listed as  $\texttt{externaldocument\{name\}}$ . In that case, we would write the above reference as

```
\externaldocument{file2}
...
...as shown in Section ref{sec:eqs} in cite{math_eqs_2020}.
```

To treat all these cases simultanesouly, DocOnce features *generalized references* which allows a reference to have different formulations, depending on whether the label is present in the document or is present in an external document. The syntax of generalized references reads

```
ref[internal][cite][external]
```

If all ref{label} references in the text internal are to labels defined in the DocOnce document, the generalized reference becomes the text internal. If one or more labels are not present in the document and latex or pdflatex is output format, the generalized reference becomes the text internal plus cite, while for all other formats the text in external is used.

The example above can now be written as the generalized reference

```
...as shown in ref[Section ref{sec:eqs}][ in cite{math_eqs_2020}][ the document "Mathematical Equations": "http://some.net/doc/matheqs.html" cite{math_eqs_2020}].
```

When \label{sec:eqs} is found in the current DocOnce document, the generalized reference becomes

```
Section ref{sec:eqs}
```

If not, and latex or pdflatex is output, the reference becomes

```
Section ref{sec:eqs}] in cite{math_eqs_2020}
```

while in all other cases the reference becomes

```
the document "Mathematical Equations":
"http://some.net/doc/matheqs.html" cite{math_eqs_2020}
```

For the reference to a label in an external document to work in the LATEX case it is required to list this document in the DocOnce file as

```
# Externaldocuments: file2
```

Several external documents can be listed with comma as delimiter:

```
# Externaldocuments: file2, file3, myfile
```

on a single line. The Externaldocuments comment leads to use of the xr package and insertion of \externaldocument{file2} in the LATEX output file. It is a good habit to place the Externaldocument comment after the title, author, and date.

Very often a reference to a chapter in a book becomes just a reference to a complete document, not a specific section, if chapters are compiled individually. For latex and pdflatex output one would then do a reference Chapter~\ref{ch:eqs} as just \cite{eqs\_doc\_2008}. That is, one would just use the cite text and not internal plus cite. This is enabled by writing the generalized reference with refch instead of ref:

```
...as shown in refch[Chapter ref{ch:eqs}][cite{eqs_doc_2008}][the document "Some Equations": "http://some.net/someeqs/"].
```

The doconce ref\_external command will read all the labels in the external documents listed in the Externaldocuments: comment and use the publish file of the current document to automatically generate substitution commands that translates ordinary references to generalized references. For example, doconce ref\_external file1 will find the reference

```
.....as shown in Section ref{sec:eqs}.
```

as a reference to an label sec:eqs defined in file2, grab the title of file2.do.txt and find the bibliographic data in the publish file and make a substitution command

```
doconce subst "Section\s+ref{sec:eqs}" "..." $files
```

where "..." is the generalized reference as shown in the examples above. In other words, with doconce ref\_external one can automatically generate generalized references between, for example, chapters in a book that exist as stand-alone documents.

#### Warning.

Generalized references to equations work well in LATEX, but not in other formats as one cannot resolve the equation number in the external document. It is then better to write different text using the FORMAT variable in Mako:

```
% if FORMAT in ("pdflatex", "latex"):
By combining ref[(ref{eqs:g1})-(ref{eqs:g4}))][ in cite{some_doc}][
dummy] we can derive the expression ...
% else:
One can from cite{some_doc} derive the expression
% endif
```

The doconce ref\_external tool generates an external text in case of references to equations that says "reference to specific **equations** (label eqs:g1 and eqs:g4) in external document "name": "link" is not recommended". One can then search for this text and make a Mako if-else rewrite as shown above.

#### Limited support.

The doconce ref\_external tool cannot correctly handle references to a range of sections like

```
Sections ref{mydoc:sec1}-ref{mydoc:sec2}
```

The automatically generated generalized references should always be manually checked and edited.

A Worked Example. Here is an example on a specific working generalized reference where the LATEX output also has a hyperlink:

```
As explained in ref[Section ref{subsec:ex}][in "Langtangen, 2012": "http://hplgit.github.io/doconce/test/demo_testdoc.html#subsec:ex" cite{DocOnce:test}][a "section": "http://hplgit.github.io/doconce/test/demo_testdoc.html#subsec:ex" in the document "A Document for Testing DocOnce": "http://hplgit.github.io/doconce/test/demo_testdoc.html" cite{DocOnce:test}], DocOnce documents may include tables.
```

With latex or pdflatex as output, this translates to

```
As explained in Section ref{subsec:ex}, DocOnce documents may include tables.
```

if the label {subsec:ex} appears in the present DocOnce source, and otherwise

```
As explained in Section ref{subsec:ex} in "Langtangen, 2012": "http://hplgit.github.io/doconce/test/demo_testdoc.html#subsec:ex" cite{DocOnce:test}, DocOnce documents may include tables.
```

The latter DocOnce code is translated to the following LATEX code:

```
As explained in Section~\ref{subsec:ex} in \href{{\http://hplgit.github.io/doconce/...}}{\Langtangen, 2012} \cite{DocOnce:test}, DocOnce documents may include tables.
```

In a format different from latex and pdflatex, the effective DocOnce text becomes

```
As explained in a "section":
"http://hplgit.github.io/doconce/test/demo_testdoc.html#subsec:ex" in the document "A Document for Testing DocOnce":
"http://hplgit.github.io/doconce/test/demo_testdoc.html"
cite{DocOnce:test}, DocOnce documents may include tables.
```

The rendered text in the current format becomes

As explained in Section ??in Langtangen, 2012 [2], DocOnce documents may include tables.

A complete "chapter" reference may look like

```
As explained in refch[Chapter ref{ch:testdoc}]["Langtangen, 2012": "http://hplgit.github.io/doconce/test/demo_testdoc.html" cite{DocOnce:test}][the document "A Document for Testing DocOnce": "http://hplgit.github.io/doconce/test/demo_testdoc.html" cite{DocOnce:test}], DocOnce documents may include tables.
```

The output now, if ch:testdoc is not a label in the document, becomes in the latex and pdflatex case

```
As explained in "Langtangen, 2012": "http://hplgit.github.io/doconce/test/demo_testdoc.html" cite{DocOnce:test}, DocOnce documents may include tables.
```

That is, the internal reference Chapter ... is omitted since it is not meaningful to refer to an external document as "Chapter". The resulting rendered text in the current format becomes

As explained in Langtangen, 2012 [2], DocOnce documents may include tables.

Note that  $\Delta T_E X$  cannot have links to local files, so a complete URL on the form http://... must be used.

#### Tip.

Use doconce ref\_external to get an overview of the external references in a file. Very often you want to rewrite the text to reduce the amount of external referencing. Remember then to compile your document before running doconce ref\_external again since the command applies the compiled files to get information (tmp\_preprocess\_\* or tmp\_mako\_\*) if you use any of the Preprocess or Mako preprocessors.

#### **2.14** Index

An index can be created for the latex, rst, and sphinx formats by the idx keyword, following a LaTeX-inspired syntax:

```
idx{some index entry}
idx{main entry!subentry}
idx{'verbatim_text' and more}
```

The exclamation mark divides a main entry and a subentry. Backquotes surround verbatim text, which is correctly transformed in a LATEX setting to

```
\index{verbatim\_text@\texttt{\rm\smaller verbatim\_text and more}}
```

Everything related to the index simply becomes invisible in plain text, Epytext, StructuredText, HTML, and wiki formats. Note: idx commands should be inserted outside paragraphs, not in between the text as this may cause some strange behaviour of reST and Sphinx formatting. As a recommended rule, index items are naturally placed right after section headings, before the text begins, while index items related to a paragraph should be placed above the paragraph one a separate line (and not in between the text or between the paragraph heading and the text body, although this works fine if LaTeX is the output format). For paragraphs with === heading, the index keywords should be placed above the heading.

The keywords in the index are automatically placed in a meta tag in html output such that search engines can make use of the them.

## 2.15 Bibliography/References

DocOnce applies the software tool Publish to handle the bibliography in a document. With Publish it is easy to import BIBTEX data and maintain a database in a clean, self-explanatory textual format. From the Publish format it is easy to go BIBTEX and reST or straightforward DocOnce typesetting (and from there to HTML, plain text, wiki formats, and so on).

Installing Publish is straightforward: just checkout the code on bitbucket. org, move to the publish directory and run sudo python setup.py install.

**Importing your data to the Publish database.** Many scientists have their bibliographic data in the BibTex format. Here we assume that you have two files, refs1.bib and refs2.bib. These can be imported to a Publish database, residing in the file papers.pub, by the commands

```
publish import refs1.bib
publish import refs2.bib
```

During import, Publish may ask you for accepting the name of new institutions or journals. Publish already have a database of journals and institutions/departments, but when you add new, you also get a file venues.list (in the current working directory) which will be used for future imports in this directory. Make sure you store publish.pub and venues.list along with your DocOnce document files (e.g., add them to your version control system).

#### Requirements to input data.

#### Notice.

Note that Publish only accepts BIBTEX files where the keys (author, title, etc.) are in lower case and where the data are enclosed in curly braces. You may need to edit your BIBTEX files to meet this demand.

The utility doconce fix\_bibtex4publish file.bib fixes several known issues with BIBTEX files such that Publish has a better chance of accepting the entries. Run this utility first, then run Publish, respond to any requirements that Publish spits out, remove papers.pub if it exists, and run the import statements again.

Although references are visible as numbers only in the output, it is recommended to have apply a nice, consistent typesetting of your keys. It is suggested to use the following scheme:

```
Langtangen_2003a  # single author
Langtangen_Pedersen_2002  # two authors
Langtangen_et_al_2002  # three or more authors
```

One can add a, b, c, and so forth if several keys feature the same authors and year.

Adding new references to the database. When you get some new BIBTEX references you simply put them in a file, say refs3.pub and run the publish import refs3.pub command to update the database. You may also consider editing the papers.pub file directly when adding new references.

**Exporting the database.** Export of everything in the database to BIBT<sub>E</sub>X is done by

```
publish export mybibtexfile.bib
```

You can easily export subsets of the database, e.g., only papers associated with a particular author (the Publish manual has details on how this is done). DocOnce will automatically export the database to BIBTEX if the output format is latex or pdflatex.

#### **Referring to publications.** We use the command

```
cite{key}
```

to refer to a publication with bibliographic key key. Here is an example: [6] discussed propagation of large destructive water waves, [5] gave an overview of numerical methods for solving the Navier-Stokes equations, while the use of Backward Kolmogorov equations for analyzing random vibrations was investigated in [4]. The book chapter [7] contains information on C++ software tools for programming multigrid methods. A real retro reference is [3] about a big FORTRAN package. Multiple references are also possible, e.g., see [6, 7].

A LATEX-style cite command with additional detailed reference is also possible,

```
cite[details]{key}
```

for example as in [5, Section 2] or [5, Equation (4.2)].

In LaTeX, the cite command is directly translated to the corresponding LaTeX version of the command with a backslash; in reST and Sphinx the citations becomes links, with the citation keys as names; in HTML the citations are numbered from 1, 2, and so forth according to their appearance, and the numbers appear as links; while in other formats the citations are simply the keys inside square brackets and the corresponding references are listed in the order they are cited.

**Specifying the Publish database.** The specification of the Publish database file in the DocOnce document is done one a line containing BIBFILE: papers.pub (you may give the database file another name and store it in another directory). The references will be inserted at the place where this command appears. Before the command you will often want to have a headline with "References", "Bibliography", or similar. Here is an example:

```
====== References ======
BIBFILE: papers.pub
```

In LATEX and PDFLATEX the papers.pub file is exported to BIBTEX format and included in the document, while in all other formats, suitable text is produced from the database.

LATEX bibliography style. The bibliography style is "plain" in LATEX output. To change this, just edit the .p.tex file. For example,

```
doconce format latex mydoc
doconce replace 'bibliographystyle{plain}' 'bibliographystyle{abbrev}' mydoc.p.tex
```

#### 2.16 Tables

A table like

time	velocity	acceleration
0.0	1.4186	-5.01
2.0	1.376512	11.919
4.0	1.1E+1	14.717624

is built up of pipe symbols and dashes:

	velocity	acceleration
0.0	1.4186   1.376512   1.1E+1	-5.01   11.919   14.717624

The pipes and column values do not need to be aligned (but why write the DocOnce source in an ugly way?). In the line below the heading, one can insert the characters c, r, or 1 to specify the alignment of the columns (centered, right, or left, respectively). Similar character can be inserted in the line above the header to algn the headings. Pipes | can also be inserted to indicate vertical rules in Latext tables (they are ignored for other formats). An example of centered headings (which is default anyway), first column left-adjusted and the others right-adjusted looks like

c	c
time	velocity   acceleration
1	r
0.0	1.4186   -5.01
1 2.0	1.376512   11.919
1 4.0	1.1E+1   14.717624

Note that not all formats offer alignment of heading or entries in tables (rst and sphinx are examples). Also note that DocOnce tables are very simple: neither entries nor headings can span several columns or rows. When that functionality is needed, one can make use of the preprocessor and if-tests on the format and insert format-specific code for tables.

The command-line option --tables2csv (to doconce format) makes Doconce dump each table to CSV format in a file table\_X.csv, where X is the table number. This feature makes it easy to load tables into spreadsheet programs for further analysis.

Data in CSV format can be transformed to DocOnce table format by the doconce csv2table utility:

Terminal> doconce csv2table somefile.csv > table.do.txt

This is a quick way of writing tables. For example, we can write a text file tmp.csv with

```
time, velocity, acceleration 0.0, 1.4186, -5.01 2.0, 1.376512, 11.919 4.0, 1.1E+1, 14.717624
```

Running doconce csv2table tmp.csv creates the table

c	cc   velocity	
c   0.0   2.0   4.0	1.4186   1.376512   1.1E+1	

If this output is redirected to a file, say mytable.do.txt, one can easily include this file by # #include "mytable.do.txt" in the DocOnce source file.

## 2.17 Exercises, Problems, Projects, and Examples

DocOnce has special support for four types of "exercises", named *exercise*, *problem*, *project*, or *example*. These are all typeset as special kind of subsections. Such subsections start with a subsection headline surrounded 5 = characters, and last up to the next headline or the end of the file. The headline itself must consists of the word *Exercise*, *Problem*, *Project*, or *Example*, followed by a colon and a title of the exercise, problem, or project. The next line(s) may contain a label and specification of the name of the result file (if the answer to the exercise is to be handed in) and a solution file. The DocOnce code looks like this:

```
==== Project: Determine the Distance to the Moon ===== label{proj:moondist} file=earth2moon.pdf solution=eart2moon_sol.do.txt

Here goes the running text of the project....
```

DocOnce will recognize the exercise, problem, project, or example *title*, the optional *label*, the optional *name of the answer file* (file=), the optional name of the *solution* file (solution=), and the *running text*. In addition, one can add subexercise environments, starting with !bsubex and ending with !esubex, on the beginning of separate lines. Within the main exercise or a subexercise, three other environments are possible: (full) solution, (short) *answer*, and *hints*. The environments have begin-end directives !bans, !eans, !bsol, !esol, !bhint, !ehint, which all must appear on the beginning of a separate line (just as !bc and !ec).

The solution environment with !bsol and !esol allows inline solution as an alternative to the solution=... directive mentioned above, which requires that the solution is in a separate file. Comment lines are inserted so that the beginning and end of answers and solutions can be identified and removed if desired.

A full exercise set-up can be sketched as follows:

```
==== Exercise: Determine the Distance to the Moon ===== label{exer:moondist} file=earth2moon.pdf

Here goes main body of text describing the exercise...
!bsubex
Subexercises are numbered a), b), etc.
!bans
Short answer to subexercise a).
!eans
!bhint
First hint to subexercise a).
!ehint
!bhint
```

```
Second hint to subexercise a).
!ehint
!esubex
!bsubex
Here goes the text for subexercise b).
!bhint
A hint for this subexercise.
!ehint
!bsol
Here goes the solution of this subexercise.
!esol
!esubex
!bremarks
At the very end of the exercise it may be appropriate to summarize
and give some perspectives. The text inside the !bremarks-!eremarks
directives is always typeset at the end of the exercise.
!eremarks
Here goes a full solution of the whole exercise.
!esol
!ec
```

A recommended rule for using the different "exercise" types goes as follows:

- Exercises are smaller problems directly related to the present chapter (e.g., with references to the text).
- Problems are sufficiently independent of the chapter's text that they make sense on their own, separated from the rest of the docoment.
- Projects are larger problems that also make sense on their own.
- Examples are exercises, problems, or projects with full solutions.

The command line options --without\_answers and --without\_solutions turn off output of answers and solutions, respectively, except for examples.

Sometimes one does not want the heading of an exercise, problem, project, or example to contain the keyword Exercise:, Problem:, Project:, or Example:. By enclosing the keyword in braces, as in

```
===== {Problem}: Find a solution to a problem =====
```

the keyword is marked for being left out of the heading, resulting in the heading "Find a solution to a problem".

The various elements of exercises are collected in a special data structure (list of dictionaries) stored in a file <code>.mydoc.exerinfo</code>, if <code>mydoc.do.txt</code> is the name of the DocOnce file. The file contains a list of dictionaries, where keys in the dictionary corresponds to elements in the exercise: filename, solution file, answer, label, list of hints, list of subexercises, closing remarks, and the main

body of text. From this data structure it is easy to generate stand-alone documents with exercises, problems, and projects with or without short answers and full solutions.

Tailored formatting of exercises in special output formats can make use of the elements in an exercise. For example, one can image web formats where the hints are displayed one by one when needed and where the result file can be uploaded. One can also think of mechanisms for downloading the solution file if the result file meets certain criteria. DocOnce does not yet generate such functionality in any output format, but this is an intended future feature to be impelemented.

For now, exercises, problems, projects, examples are typeset as ordinary DocOnce sections (this is the most general approach that will work for many formats). One must therefore refer to an exercise, problem, project, or example by its label, which normally will translate to the section number (in LaTeX, for instance) or a link to the title of the section. The *title* is typeset without any leading *Exercise:*, *Problem:*, or *Project:* word, so that references like

```
see Problem ref{...}
```

works well in all formats (i.e., no double Problem Problem appears).

**Remark.** Examples are *not* typeset similarly to exercises unless one adds the command-line option —examples\_as\_exercises. That is, without this option, any heading and starting with Example: makes DocOnce treat the forthcoming text as ordinary text without any interpretation of exercise-style instructions. With the command-line option —examples\_as\_exercises, one can use the !bsubex and !bsol commands to indicate a subproblem and a solution. In this way, the typesetting of the example looks like an exercise equipped with a solution.

**List of Exercises, Problems, and Projects.** DocOnce also supports listing all exercises, problems, and projects with corresponding page numbers. By default, no such listing is enabled. When running doconce format pdflatex (or latex) there is a command-line option --list\_of\_exercises that can be set to

- toc: include exercises, problems, and projects as part of the table of contents
- loe: make a separate list of exercises, problems, and projects, which appears after the table of contents

There is a special LaTEX environment for the exercises that one can use to further tailor the appearance of exercises (given that one edits the .tex file, preferably by a script after each doconce format command).

### 2.18 Blocks of Verbatim Computer Code

Blocks of computer code, to be typeset verbatim, must appear inside a "begin code" !bc keyword and an "end code" !ec keyword. Both keywords must be on a single line and *start at the beginning of the line*. Before such a code block there must be a plain sentence (at least if successful transformation to reST and ASCII-type formats is desired). For example, a code block cannot come directly after a section/paragraph heading or a table.

Here is a plain code block:

```
!bc
% Could be a comment line in some file
% And some data
1.003 1.025
2.204 1.730
3.001 1.198
!ec
which gets rendered as
```

```
% Could be a comment line in some file
% And some data
1.003 1.025
2.204 1.730
3.001 1.198
```

There may be an argument after the !bc tag to specify a certain environment (for ptex2tex, doconce ptex2tex, or Sphinx) for typesetting the verbatim code. For instance, !bc dat corresponds to the data file environment and !bc cod is typically used for a code snippet. There are some predefined environments explained below. If there is no argument specifying the environment, one assumes some plain verbatim typesetting (for ptex2tex this means the ccq environment, which is defined in the config file .ptex2tex.cfg, while for Sphinx it defaults to the python environment).

By default, pro is used for complete programs in Python, cod is for a code snippet in Python, while xcod and xpro implies computer language specific typesetting where x can be f for Fortran, c for C, cpp for C++, sh for Unix shells, pl for Perl, m for Matlab, cy for Cython, r for Ruby, and py for Python. The argument sys means by default console for Sphinx and CodeTerminal (ptex2tex environent) for LaTeX. Other specifications are dat for a data file or print out, latex for LaTeX', js for JavaScript, html for HTML, pyshell for plain interactive Python shell sessions, and ipy for interactive IPython sessions (the latter does not introduce any environment in sphinx output, as interactive sessions are automatically recognized and handled in that format). (All these definitions of the arguments after !bc can be redefined in the .ptex2tex.cfg configuration file for ptex2tex/LaTeX and in the sphinx code-blocks comments for Sphinx, see below.)

Since the config file for ptex2tex and command-line arguments for the alternative doconce ptex2tex program can define what some environments map onto with respect to typesetting, a similar possibility is supported for Sphinx as

well. The argument after !bc is in case of Sphinx output mapped onto a valid Pygments language for typesetting of the verbatim block by Pygments. This mapping takes place in an optional comment to be inserted in the DocOnce source file, e.g.,

# sphinx code-blocks: pycod=python cod=fortran cppcod=c++ sys=console

Here, three arguments are defined: !bc pycod maps on to the Pygments style python for Python code, !bc cod maps on to the Pygments style fortran for Fortran code, !bc cppcod maps on to the Pygments style c++ for C++ code, and !bc sys maps on to the Pygments console style for terminal sessions. The same arguments would be defined in .ptex2tex.cfg or on the command line for doconce ptex2tex for how to typeset the blocks in LaTeX using various verbatim styles (Pygments can also be used in a LaTeX context).

#### Tip.

The enclosing !ec tag of verbatim computer code blocks must be followed by a newline. A common error in list environments is to forget to indent the plain text surrounding the code blocks. In general, we recommend to use paragraph headings instead of list items in combination with code blocks (it usually looks better, and some common errors are naturally avoided).

Here is a verbatim code block with Python code (pycod style):

```
!bc pycod
def f(x, y):
    return x + y

# Main program
from math import pi
print 'Testing f:', f(pi, 0)
lec
```

The typeset result of this block becomes

```
def f(x, y):
    return x + y

# Main program
from math import pi
print 'Testing f:', f(pi, 0)
```

And here is a C++ code snippet (cppcod style):

```
!bc cppcod
void myfunc(double* x, const double& myarr) {
    for (int i = 1; i < myarr.size(); i++) {
        myarr[i] = myarr[i] - x[i]*myarr[i-1]
    }
}
!ec</pre>
```

with the rendered result

```
void myfunc(double* x, const double& myarr) {
   for (int i = 1; i < myarr.size(); i++) {
      myarr[i] = myarr[i] - x[i]*myarr[i-1]
   }
}</pre>
```

Computer code can be copied directly from a file, if desired. The syntax is then

```
@@@CODE myfile.f
@@@CODE myfile.f fromto: subroutine\s+test@^C\s{5}END1
```

The first line implies that all lines in the file myfile.f are copied into a verbatim block, typset in a !bc Xpro environment, where X is the extension of the filename, here f (i.e., the environment becomes !bc fpro and will typically lead to some Fortran-style formatting). The second line has a fromto: directive, which implies copying code between two lines in the code, typset within a ¡bc Xcod' environment (again, X is the filename extension, implying the type of file). Note that the pro and cod arguments are only used for LaTeX and Sphinx output, all other formats will have the code typeset within a plain !bc environment.) Two regular expressions, separated by the @ sign, define the "from" and "to" lines. The "from" line is included in the verbatim block, while the "to" line is not. In the example above, we copy code from the line matching subroutine test (with as many blanks as desired between the two words) and the line matching C END1 (C followed by 5 blanks and then the text END1). The final line with the "to" text is not included in the verbatim block.

One can also specify the code environment explictly rather than relying on the file extension:

```
@@@CODE somefile.py envir=X fromto: def myfunc@def yourfunc
```

This is the same as writing !bc X and copying parts of the somefile.py text into the DocOnce source file, followed by !ec. For example, if the lines between the myfunc and yourfunc functions actually work as a complete Python program, one could specify envir=pypro to indicate that it is a complete program and not a snippet (i.e., pycod would be the default without envir=pypro in this case).

Let us copy a whole file (the first line above):

```
program testme
call test()
return
```

Let us then copy just a piece in the middle as indicated by the fromto: directive above:

```
subroutine test()
integer i
real*8 r
r = 0
do i = 1, i
    r = r + i
end do
return
```

Note that the "to" line is not copied into the DocOnce file, but the "from" line is. Sometimes it is convenient to also neglect the "from" line, a feature that is allowed by replacing fromto: by from-to ("from with minus"). This allows for copying very similar code segments throughout a file, while still distinguishing between them. Copying the second set of parameters from the text

```
# --- Start Example 1 ---
c = -1
A = 2
p0 = 4
simulate_and_plot(c, A, p0)
# --- End Example 1 ---
# --- Start Example 2 ---
c = -1
A = 1
p0 = 0
simulate_and_plot(c, A, p0)
# --- End Example 2 ---
is easy with
from-to: Start Example 2@End Example 2
```

With only fromto: this would be impossible.

Remark for those familiar with ptex2tex: The from-to syntax is slightly different from that used in ptex2tex. When transforming DocOnce to LaTeX, one first transforms the document to a .p.tex file to be treated by ptex2tex. However, the @@@CODE line is interpreted by DocOnce and replaced by the mentioned pro or cod environment which are defined in the ptex2tex configuration file.

Remark for those familiar with the listings package in LTEX: the listing package can copy code from files, but snippets must be specified through exact line numbers. The @@@CODE directive above works with regular expressions which are much less sensitive to edits of the source code file than the line numbers. Moreover, copy of code from file works in Dococe across formats (HTML, Sphinx, Markdown, etc.).

# 2.19 LATEX Blocks of Mathematical Text

Blocks of mathematical text are like computer code blocks, but the opening tag is !bt (begin TeX) and the closing tag is !et. It is important that !bt and !et appear on the beginning of the line and followed by a newline.

```
!bt
\begin{align}
{\partial u\over\partial t} &= \nabla^2 u + f, label{myeq1}\\
{\partial v\over\partial t} &= \nabla\cdot(q(u)\nabla v) + g. label{myeq2}\end{align}
```

The support of LATEX mathematics varies among the formats:

- Output in Latex and pdflatex formats) has of course full support of all LateX mathematics, of course.
- The html format supports single equations and multiple equations via the align environment, also with labels.
- Markdown (pandoc format) allows single equations and inline mathematics.
- MediaWiki (mwiki format) does not enable labels in equations and hence equations cannot be referred to.

The main conclusion is that for output beyond LATEX (latex and pdflatex formats), stick to simple \[ and \] or equation and align or align\* environments, and avoid referring to equations in MediaWikis.

Going from DocOnce to MS Word is most easily done by outputting in the latex format and then using the Pandoc program to translate from LATEX to MS Word (note that only a subset of LATEX will be translated correctly).

If the document targets formats with and without support of LaTeX mathematics, one can use the preprocessor to typeset the mathematics in two versions. After #if FORMAT in ("latex", "pdflatex", "html", "sphinx", "mwiki", one places LaTeX mathematics, and after #else one can write inline mathematics in a way that looks nice in plain text and wiki formats without support for mathematical typesetting. Such branching can be used with make if-else statements alternatively:

```
"pandoc")
```

**Mathematics for PowerPoint/OpenOffice.** If you have LATEX mathematics written in DocOnce, it is fairly easy to generate PNG images of all mathematical formulas and equations for use with PowerPoint or OpenOffice presentations.

- 1. Make a Sphinx version of the DocOnce file.
- 2. Go to the Sphinx directory and load the conf.py file into a browser.
- 3. Search for "math" and comment out the 'sphinx.ext.mathjax' (enabled by default) and 'matplotlib.sphinxext.mathmpl' (disabled by default) lines, and uncomment the 'sphinx.ext.pngmath' package. This is the package that generates small PNG pictures of the mathematics.
- 4. Uncomment the line with pngmath\_dvipng\_args = and set the PNG resolution to -D 200 when the purpose is to generate mathematics pictures for slides.
- 5. Run make html.
- 6. Look at the HTML source file in the \_build/html directory: all mathematics are in img tags with src= pointing to a PNG file and alt= pointing to the LaTEX source for the formula in question. This makes it very easy to find the PNG file that corresponding to a particular mathematical expression.

# 2.20 Macros (Newcommands)

DocOnce supports a type of macros via a LaTeX-style newcommand construction. The newcommands defined in a file with name newcommand\_replace.tex are expanded when DocOnce is filtered to other formats, except for LaTeX (since LaTeX performs the expansion itself). Newcommands in files with names newcommands.tex and newcommands\_keep.tex are kept unaltered when DocOnce text is filtered to other formats, except for the Sphinx format. Since Sphinx understands LaTeX math, but not newcommands if the Sphinx output is HTML, it makes most sense to expand all newcommands. Normally, a user will put all newcommands that appear in math blocks surrounded by !bt and !et in newcommands\_keep.tex to keep them unchanged, at least if they contribute to make the raw LaTeX math text easier to read in the formats that cannot render LaTeX. Newcommands used elsewhere throughout the text will usually be placed in newcommands\_replace.tex and expanded by DocOnce. The definitions of newcommands in the newcommands\*.tex files must appear on a single line (multi-line newcommands are too hard to parse with regular expressions).

#### 2.21 Admonitions

DocOnce offers strong support for admonition environments, such as warning boxes, notification boxes, question boxes, and summary boxes. The boxes

normally have an icon, a heading, and may also have a background color. A special box, the block, has never any icon and can be used when an icon would be disturbing or misleading.

The following admonition environments are available: block, warning, notice, question, and summary. The box is defined by begin and end tags such as !bnotice and !enotice. The title of the box is fully customizable.

Here are a few examples:

```
!bwarning
Here is a warning!
!ewarning
!bnotice Hint
This is a hint.
!enotice
!bblock This is a block.
A block has never any icon. It is mostly used in slides
and gives the text block effect in LaTeX beamer.
!eblock
!bnotice Going deeper
This is text meant to provide more details. The box has the
layout of the notice box, but a custom title, here *Going deeper*.
!enotice
Finally some summary:
!bsummary
The main message is to utilize the admonition styles for
marking different parts of the text
!esummary
```

The above DocOnce code is in the present format rendered as

# Warning. Here is a warning!

Hint.
This is a hint.

This is a block.

A block has never any icon.

#### Going deeper.

This is text meant to provide more details. The box has the layout of the notice box, but a custom title, here *Going deeper*.

Finally some summary:

#### Summary.

The main message is to utilize the admonition styles for marking different parts of the text

The layout of admonitions depend on the format. In rst and sphinx one applies the native admonitions, but in sphinx the automake\_sphinx.py script manipulates the HTML file to set a gray background for all admonitions. In html one has a command-line argument --html\_admon that can be set to different styles:

- --html\_admon=gray for icons with gray background and small font,
- --html\_admon=yellow and --html\_admon=apricot are similar, but the icons and colors are different,
- --html\_admon=colors has quite bright colors as backgrounds for the different admonitions,
- --html\_admon=lyx gives a white background and small icons,
- --html\_admon=paragraph results in a simple paragraph with the admon title as heading.

The options --html\_admon\_bg\_color=... and --html\_admon\_bd\_color=... can be used to override the default background and boundary frame colors of the admon styles (respectively). These options have only effect for the apricot, yellow, and gray styles.

Some recommended combinations for admonitions in HTML are

- --html\_style=bloodish, --html\_admon=gray, --pygments\_html\_style=none Or --pygments\_html\_style=default
- --html\_style=boots\*, --html\_admon=bootstrap\_alert Or --html\_admon=bootstrap\_panel

The --html\_admon style has no effect for the vagrant style template in bundled/html\_styles/style\_vagrant. In latex and pdflatex, the type of admonition style is set by the command-line option --latex\_admon=. Several values are available:

- paragraph is the simplest type of admonition and typeset as plain text with an optional paragraph heading.
- colors1 (inspired by the NumPy User Guide) applies different colors for the different admons with an embedded icon.
- colors2 is like colors1 but the text is wrapped around the icon.
- mdfbox is the default and gives rounded (gray) boxes with a potential title and no icon (using the very flexible mdframed packaged in LATEX)
- graybox2 has square corners, gray background, and is narrower than mdfbox. One special feature of graybox2 is the summary admon, which has a different look with horizontal rules only, and for A4 format, the summary box is half of the text width and wrapped with running text around (if it does not contain verbatim text, in that case the standard graybox2 style is used). This small summary box is effective in proposals to disperse small paragraphs of key points around.
- grayicon has icons and a light gray background.
- yellowicon has icons and a light yellow background.

There is also an option <code>--latex\_admon\_color=...</code> that can be used to override the default color. Values are either saturated colors like <code>gray!10</code> or an RGB tuple <code>0.95,0.91,0.97</code>. The chosen color replaces all default colors for all admon styles except <code>paragraph</code>. For example, an oval white box is produced by <code>--latex\_admon=mdfbox</code> and <code>--latex\_admon\_color=white</code>.

**The Box Environment.** A plain box without any title or colored background, just a rectangular frame, is also available. The syntax goes like

```
!bbox
!bt
\[ \nabla\cdot\bm{u} = 0 \quad\hbox{(mass balance)}\]
!et
!ebox
```

resulting in

```
abla \cdot oldsymbol{u} = 0 (mass balance)
```

Admonitions are usually used to typeset something that should be distinct from the running text: a warning, a remark, a notification, a question, a summary, while a box is often used to highlight a key finding in the running text: an equation, a hypothesis, a theorem, a rule, or a conclusion.

# 2.22 Typesetting of Algorithms

DocOnce has no support for typesetting of algorithms, while LaTEX has quite sophisticated support. LaTEX writes need some strategy with dealing with algorithms in DocOnce: they should be sophisticated when the output is in LaTEX and they should be readable when the output is in other formats that have no native support for algorithms.

The simplest solution is to use the preprocessor (Preprocess or Mako) to allow different solutions for different formats. There are basically three typesetting styles available: LATEX algorithm environments, pseudo code in a verbatim block, and a nested list. A preprocessor variable, say ALG can be used to select the typesetting. With Mako we can write

```
% if ALG == 'latex':
# Write native LaTeX code for the algorithm
% elif ALG == 'code':
# Write pseudo code in (e.g.) a python-like style
!bc pycode
if feature in element:
...
!ec
% elif ALG == 'list':
# Use lists to express the algorithm
    o if this element has the feature:
        * ...
% endif
```

# 2.23 Preprocessing Steps

DocOnce allows preprocessor commands for, e.g., including files, leaving out text, or inserting special text depending on the format. Two preprocessors are supported: preprocess (http://code.google.com/p/preprocess) and mako (http://www.makotemplates.org/). The former allows include and if-else statements much like the well-known preprocessor in C and C++ (but it does not allow sophisticated macro substitutions). The latter preprocessor is a very powerful template system. With Mako you can automatically generate various type of text and steer the generation through Python code embedded in the DocOnce document. An arbitrary set of name=value command-line arguments (at the end of the command line) automatically define Mako variables that are substituted in the document.

DocOnce will detect if preprocess or Mako commands are used and run the relevant preprocessor prior to translating the DocOnce source to a specific format.

The preprocess and make programs always have the variable FORMAT defined as the desired output format of DocOnce (html, latex, plain, rst, sphinx, epydoc, st). It is then easy to test on the value of FORMAT and take different actions for different formats. Below is an example:

First some math:

```
\begin{align}
x \&= 3
label{x:eq1}\\\
y &= 5
label{y:eq1}
\end{align}
!et.
Let us now reason about this.
# Sphinx cannot refer to labels in align environments
# #if FORMAT in ("latex", "pdflatex", "html")
From (\ref{x:eq})-(\ref{y:eq1}) we get that # #elif FORMAT == "sphinx"
From
[x = 3]
!et
and
!bt
\[ y= 5 \]
!et
it follows that
# #else
From the above equations it follows that
# #endif
x+y is 8.
```

A variable DEVICE is also defined. It equals screen by default, but the command-line argument --device=paper can set DEVICE to paper (or another value). Testing on DEVICE inside the document makes it possible to test if the output is on paper media, a sreen, or a particular device.

Other user-defined variables for the preprocessor can be set at the command line as explained in Section 5.

More advanced use of make can include Python code that may automate the writing of parts of the document.

# 2.24 Splitting Documents into Smaller Pieces

Long documents are conveniently split into smaller DocOnce files. However, there must be a master document including all the pieces, otherwise references to sections and the index will not work properly. The master document is preferably a file just containing a set of preprocessor include statements of the form #include "file.do.txt". The preprocessor will put together all the pieces so that DocOnce sees a long file with the complete text.

For web documents it is often desired to split long pages into shorter ones. This is done by the DocOnce command !split placed at the beginning of a line. The !split commands works with output in html, rst, sphinx, latex, and pdflatex. The !split command are normally placed before section headings. It is very actively used when writing slides with DocOnce. The doconce format command does not recognize !split instructions: one needs to run doconce split\_\* as a postprocess, where the \* means html, rst, or beamer.

#### **HTML.** Splitting an HTML document is done by

```
Terminal> doconce format html mydoc
Terminal> doconce split_html mydoc
```

The mydoc.html document created by the first command is replaced by a new HTML file, representing the first part of the document, after the second command. The various files that constitute the parts of the document are listed after the split\_html command. The files have names mydoc.html, .\_mydoc000.html (equal to mydoc.html), .\_mydoc001.html, .\_mydoc002.html, and so on. Recall that all the parts are needed if the HTML document is to be moved to another location (you can always check .mydoc\_html\_file\_collection for a list of all the files that are needed to display this HTML document).

MathJax cannot refer to equations defined in other HTML files, but the doconce split\_html fixes this problem. Note, however, that running doconce split\_html leads to another equation numbering than in the original HTML document. In the latter, we use AMS equation numbering, which means that the standard Later conventions are followed, while in the splitted document only the subset of equations with labels are given numbers.

The HTML documents have by default very simple navigation buttons for the previous and next document. These buttons are customizable:

Terminal> doconce split\_html mydoc.html --nav\_buttons=X

where X can be text (pure text "Previous" and "Next", no buttons), gray1, gray2, bigblue, blue, or green as shown here, respectively (from left to right):



However, if --html\_theme= is set any theme starting with bootstrap or bootswatch, the navigation buttons are ignored, and Bootstrap-style buttons are used.

**reStructuredText and Sphinx.** Here is a typical split of a large Sphinx document mydoc.rst into smaller pieces:

The doconce format sphinx mydoc command is needed to produce mydoc.rst, which is the starting point for the doconce split\_rst command. The various files that constitute the complete Sphinx document are mydoc.rst, .\_mydoc000.rst, .\_mydoc001.rst, .\_mydoc002.rst, and so on. The automake\_sphinx.py script ensures that the Sphinx document is compiled correctly. If all links to local files

are in a \_static directory, the whole Sphinx document exists in a complete version in the compiled directory (usually sphinx-rootdir/\_build/html) and can easily be moved around.

LATEX Beamer slides generated from DocOnce source also apply !split to indicate the start of individual slides. However, the split is performed by the doconce slides\_beamer command and does not result in individual files like split\_rst and split\_html do.

#### 2.25 Header and Footer

Some formats use a header and footer in the document. LateX and HTML are two examples of such formats. When the document is to be included in another document (which is often the case with DocOnce-based documents), the header and footer are not wanted, while these are needed (at least in a LateX context) if the document is stand-alone. We have introduced the convention that if TITLE: is found at the beginning of the line (i.e., the document has a title), the header and footer are included, otherwise not.

# 2.26 Missing Features

DocOnce does not aim to support sophisticated typesetting, simply because sophisticated typesetting usually depend quite strongly on the particular output format chosen. When a particular feature needed is not supported by DocOnce, it is recommended to hardcode that feature for a particular format and use the if-else construction of the preprocessor. For example, if a sophisticated table is desired in LATEX output, do something like

```
# #if FORMAT in ("latex", "pdflatex")
# insert native LaTeX code for fancy table
# #else
# insert a DocOnce-formatted "inline" table
# #endif
```

Similarly, if certain adjustments are needed, like pagebreaks in LaTeX, hard-code that in the DocOnce format (and recall that this is really LaTeX dependent - pagebreaks are not relevant HTML formats).

Instead of inserting special code in the DocOnce document, one can alternatively script editing of the output from DocOnce. That is, we develop a Python or Bash script that runs the translation of a DocOnce document to a ready docoment in another format. Inside this script, we may edit and fine-tune the output from DocOnce.

# 2.27 Guidelines for LATEX Writers

Any LaTeX syntax is accepted, but if output in the sphinx, pandoc, mwiki, html, or ipynb formats is also important, one should follow these rules:

- AMS Lagrange TeX mathematics is supported, also for the html, sphinx, and ipynb formats.
- Only five equation environments can be used: \[ ... \], equation\*, equation, align\*, and align.
- Newcommands in mathematical formulas are allowed, but not in the running text. Newcommands must be defined in files with names newcommands\*.tex.
- MediaWiki (mwiki) does not support references to equations.

(DocOnce performs extensions to sphinx and other formats such that labels in align environments work well.)

#### Notice.

Later X supports lots of fancy formatting, for example, multiple plots in the same figure (subfigure package), margin notes, multi-page tables, example and algorithm environments, code listings in figures, etc. Other output formats, such as sphinx and html, make it necessary for DocOnce to only utilze very standard Later X-style writing. Correspondingly, the writer must stick to simple constructs and avoid, for instance,

- more than one plot per figure (combine them into one using doconce combine\_images),
- floating tables with captions (tables are inlined and explained in the text),
- example environments (use subsection or subsubsection instead),
- algorithms environments (see *Typesetting of Algorithms* in this manual).
- code listings in figures (use just inline blocks).

However, one can use preprocessor if-tests on the format (typically #if FORN AT in ("latex", "pdflatex")) to include special code for latex and pdflatex output and more straightforward typesetting for other formats. In this way, one can also allow advanced Latex and fine tuning of the resulting PDF document. More tuning can be done by automatic editing of the .tex produced by DocOnce.

One can also use the Mako preprocessor to extend DocOnce functionality and implement, e.g., theorem or example environments (see Section 4.5) or tables with figures (see Section 4.4). Many other extensions are also possible.

## 2.28 Emacs DocOnce Formatter

The file .doconce-mode.el in the DocOnce source distribution gives a "DocOnce Editing Mode" in Emacs.

Here is how to get the DocOnce Editing Mode in Emacs: Download .doconce-mode.el and save it in your home directory, then add these lines to ~/.emacs:

```
(load-file "~/.doconce-mode.el")
```

Emacs will now recognize files with extension .do.txt and enter the DocOnce Editing Mode.

The major advantage with the DocOnce Editing Mode in Emacs is that many keyboard shortcuts are defined:

Emacs key	Action
Ctrl+c f	figure
Ctrl+c v	movie/video
Ctrl+c h1	heading level 1 (section/h1)
Ctrl+c h2	heading level 2 (subsection/h2)
Ctrl+c h3	heading level 2 (subsection/h3)
Ctrl+c hp	heading for paragraph
Ctrl+c me	math environment: !bt equation !et
Ctrl+c ma	math environment: !bt align !et
Ctrl+c ce	code environment: !bc code !ec
Ctrl+c cf	code from file: @@@CODE
Ctrl+c table2	table with 2 columns
Ctrl+c table3	table with 3 columns
Ctrl+c table4	table with 4 columns
Ctrl+c exer	exercise outline
Ctrl+c slide	slide outline
Ctrl+c help	print this table

Typing Ctrl+c help prints the above table in Emacs. Try out the different short-cuts and see how handy they are in learning DocOnce and saving much typing!

# 3 Writing Slides

It is a fast procedure to make slides from large amounts of DocOnce text, in particular for condensing running material for teaching or just providing the slide set as an overview or study guide. The slides can either be ordinary, separate slides - or just a document with much briefer text and emphasis on bullet lists, figures, mathematical formulas, admonitions, and little text.

Points to consider (just sketches...):

 Only some pygments style are suited for a particular reveal.js/deck.js theme

- Only some admon styles are appropriate
- Admon styles are erased in reveal
- Use --keep\_pygments\_html\_bg to avoid big changes in background color for code
- Can use PowerPoint or Google Presenter to design a slide and then save as image (GP directly, PP via Save as Picture or Save As + choosing image type, and then all or present slide only), then import image in doconce. LATEX formulas in image are easiest created as PNGs using http://www.codecogs.com/latex/eqneditor.php.
- From doconce to PP or GP: use the latexslides script to convert to PDFs and odp format, convert odp to ppt interactively or with unoconv.

#### 3.1 Overview

Basically, DocOnce slides are ordinary DocOnce text with !split inserted before each slide. Nevertheless, contents of slide differ considerably from ordinary running text. Some guidelines on the elements within each slide are necessary to produce effective slide sets:

- Use a subsection heading as slide heading (5 =).
- Limit the amount of running text (as always).
- Limit the amount of material so it fits within a slide (inspect slides visually to move or delete content - just an extra !split and a new heading is enough to make a new slide).
- Use the slidecell environment (see below) to create a grid of slide cells (makes it easy to move figures and bullet lists or text around).
- Adjust the size of figures (width parameter for HTML, frac parameter for LATEX Beamer) so they become effective on the slide.

DocOnce can generate two types of slides: HTML5+CSS3 type of slides to be presented through a web browser, and classical LATEX Beamer slides.

## 3.2 HTML5 Slides

Not yet written...

Just a very preliminary sketch of commands:

#### Potential Problems.

- Some newer Firefox does not show rounded corners of the admonition boxes, e.g., notice and warning (tested on Ubuntu)
- DocOnce performs some adjustments of the spacing around equations.
   More edits (automate with doconce subst) might be needed.

# 3.3 LATEX Beamer Slides

Not yet written...

**Themes.** Four themes come with DocOnce: X\_Y, where X is blue or red (the main color of the slides) and Y is plain for simple layout and shadow for shadowed boxes and more visual structure in the slides.

# 4 Mako Programming

The Mako templating engine is used as preprocessor for DocOnce documents, but the Preprocess is run prior to Mako and is recommended for including other files via # #include "filename". Preprocess is also sufficient for if-else tests to steer which parts of the text that are to be compiled. For more demanding tasks, use Mako, which resembles a real programming language.

#### Warning.

Unfortunately, the combination of Mako and LTEX mathematics may lead to problems because Mako applies syntax like \${var}\$ to extract variables or call functions, while LTEX mathematics sometimes applies the same syntax, e.g., \${\cal 0}(\Delta x^2)\$ which looks like a Mako function call. This problem can give rise to strange error messages from Mako, usually that a variable is not defined. The solution is to avoid such Makostyle syntax when writing LTEX mathematics, e.g., by defining newcommands if it is otherwise problematic.

#### 4.1 The Basics of Mako

Just a preliminary sketch of some Mako code (next example is better!):

```
# Define variables
<%
mycounter = 1
mydict = {}
%</pre>
```

```
\mbox{\#} Assume MYVAR is given on the command line as MYVAR=mytext (e.g.) % if MYVAR is not UNDEFINED:
The value of MYVAR is ${MYVAR}.
% endif
## Manipulation of variables
mycounter += 1
% if MYVAR in (2,4,6):
MYVAR is even integer in [2,6].
% elif MYVAR > 1000000:
MYVAR is big.
% else:
MYVAR=${MYVAR}, mycounter=${mycounter}.
% endif
# Function
<%
# Define Python function: FORMAT and DEVICE
# are always defined
def link(filename):
    url = "https://github.com/some/path/to/' + filename + '"'
    if DEVICE == 'screen':
         # make link to url
return '"filename":' + url
    elif DEVICE == 'paper':
         # write URL explicit on paper
return 'URL:' + url
%>
<%doc>
This
is
block
comment in Mako.
<%doc/>
```

# 4.2 Debugging Python code in Mako

Although it is easy to put in Python code inside <% and %> directives, it may quickly become a nightmare to debug the Python code. If Mako reports a syntax error in the Python code, it is recommended to debug the Python code outside of Mako. That is, copy all Python code to a separate file and run it as standard Python code. This will give much more precise (and familiar) error messages.

A setup that is easy for developing and maintaining the Python code used in Mako goes as follows. Different pieces of Python code in Mako is placed in separate files and included via the Preprocess preprocessor in the DocOnce document. For instance,

```
## Define Mako variables and basic functions
<%
# #include "variables.py"</pre>
```

```
# #include "basic_functions.py"
%>
Some DocOnce text ....
## Define more functions
<%
# #include "more_functions.py"
%>
```

Here, we have separated the Python code in Mako into three files: variables.py, basic\_functions.py, and more\_functions.py. With the # #include statements, Mako sees the .py files inserted as text inside the <% and %> directives (check the file tmp\_preprocess\_\_mydoc.do.txt for the result of running Preprocess and the complete DocOnce file that Mako sees).

To debug the .py files, collect all their text in one single .py file and run it. The recommended way is to use Preprocess to copy all the text in the .py files into a single file, named (say) all\_code.p.py:

```
# --*-- coding: utf-8 -*-
# (Encoding info cannot be in make code, but required here
# in Python code if we use non-ascii characters)

# #include "variables.py"
# #include "basic_functions.py"
# #include "more_functions.py"

# Test the code included above
print some_func(1, 'arg')
```

Note that after including the .py files one can insert calls to functions, print variables, etc., as desired to debug and experiment.

Run Preprocess on all\_code.p.py to produce an ordinary Python file all\_code.py and run that code to test and debug. A little shell script all\_code.sh doing both these things is convenient:

```
#!/bin/sh
preprocess all_code.p.py > all_code.py
python all_code.py
```

Especially when the Python code in Mako grows and becomes non-trivial, this setup for having the code external to the DocOnce document is very effective.

One can use Mako to extend the DocOnce syntax. Three examples are given next.

# 4.3 Example: Executing Python and using SymPy Objects in LATEX

Here is an example where we want to illustrate how to calculate a double integral. All mathematics is to be done in SymPy, and results are supposed to be embedded in the document's text. We can include Python code to be executed, anywhere in the document, and a variable var in the Python code is reached by

\$\{\var\}. Here, we run SymPy and use SymPy's LATEX converter to make LATEX code out of computed SymPy expressions:

```
# Execute Python code

//
import sympy as sm
x, y, a = sm.symbols('x y a')
f = a*x + sm.sin(y)
step1 = sm.Integral(f, x, y)
step2 = sm.Integral(sm.Integralf, x).doit(), y)
step3 = step2.doit()
%>

# Make use of results in the above block when writing LaTeX math
!bt
\begin{align*}
${sm.latex(step1)} &= ${sm.latex(step2)}\\
&= ${sm.latex(step3)}\
end{align*}
!et
```

The result of the LATEX block above, after Mako is run, becomes

```
\begin{align*}
\iint a x + y^{2} \sin{\left (y \right )}\, dx\, dy &=
\int \frac{a x^{2}}{2} + x y^{2} \sin{\left (y \right )}\, dy\\
&= \frac{a y}{2} x^{2} + x \left(- y^{2} \cos{\left (y \right )} +
2 y \sin{\left (y \right )} + 2 \cos{\left (y \right )}\\
\end{align*}
```

Debugging Python code in Mako is less convenient than debugging Python files directly, so one may prefer to just include the Python code that Mako is supposed to run by

```
<%
# #include "src/ex1.py"
%>
```

This is the recommended way to make use of SymPy to automate the mathematical derivations: first develop and test the SymPy code files, include the files in the document inside Mako's Python code environment.

**Remark.** Executing Python code inside the DocOnce document is closely related to *literate programming*. Tools supporting creating a document while running a Python programming cover Pweave, Python literate, PyWebTool, Antiweb, Literate Programming in Python, Pyreport, and also IPython notebooks.

# 4.4 Example: Extending Tables to Handle Figures

DocOnce tables cannot contain figures, since figures must appear on a single line with the special syntax FIGURE: [file, ...]. What if you want a table of thumbnail figures with hyperlinks?

**Solution for LATEX and HTML.** We first restrict the attention to LATEX and HTML. In those cases we can create an ordinary table and insert a call to a Mako function in cells in the table to return the proper LATEX or HTML code for displaying a figure.

The Mako call syntax in a cell is decided to be \$\{tfig('080')\}, which insert the figure mov/wave\_frames/frame\_0080.png. A 3x3 table of figures then look like

We do not want a heading, just a grid of figures, but a heading is required in DocOnce figures, so the solution is to provide empty column names in the heading. This solution gives acceptable results in HTML and LATEX.

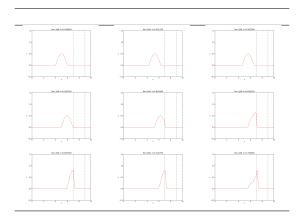
The Mako function can be a plain Python function:

Note that for other formats than LaTEX and HTML we just return a link to the figure.

#### Notice.

A corresponding solution for Sphinx will not work because inline figures in Sphinx has a syntax with pipe symbols ('|') that interfer with the column separator in tables in DocOnce. Returning the HTML code in case of Sphinx will just display that HTML code in the cells, not the rendered HTML code. A solution that includes Sphinx is provided later.

The resulting table is displayed below.



**Generating the Entire Table.** The specific structure of this table also suggest using Mako to generate the entire table:

```
def generate_table(start, end, step, no_of_columns):
   # Heading
   text = '
   horizontal_rule = '|----|\n'
   text += horizontal_rule
    text += '| '*no_of_columns + '|\n'
   text += horizontal_rule
   fig_counter = 0
    for counter in range(start, end+1, step):
        fig_counter += 1
        text += '| ' + tfig('%03d' % counter) + ' '
        if fig_counter % no_of_columns == 0:
           text += '|\n'
    text += horizontal_rule
   return text
%>
${generate_table(80, 120, 5, 3)}
```

The result is a table like the one above, except that the horizontal rules and the heading are very short lines (but this is legal syntax - it just does not look appealing in the DocOnce source).

**Generating the Entire Table in a Native Format via Mako.** To allow Sphinx and other formats, it is best to generate the entire table. This can be done either by a Mako function or as a post process after DocOnce has created the output file. We show the Mako solution here.

```
Tip.
```

A pure Python function Mako is easiest to develop in a seperate Python program file because Python has better error messages than Mako. Also, Python has debugging facilities that make the development of the code much more efficient than writing in Mako. When the Python function works, it can be inserted in the DocOnce file inside the Mako directives <% and %>.

The Mako function for generating the entire table in LATEX, HTML, and Sphinx format goes as follows.

```
def generate_native_table(start, end, step, no_of_columns):
   text = '
   # Heading
   if FORMAT in ("latex", "pdflatex"):
       text += r"""\begin{quote}\begin{tabular}{%s}
\hline
11
\hline
""" % ('c'*no_of_columns)
   elif FORMAT in ("sphinx", "rst", "html"):
       text += '\n'
   fig_counter = 0
   latex_columns = []
   for counter in range(start, end+1, step):
       fig_counter += 1
       if FORMAT in ("latex", "pdflatex"):
            latex_columns.append(r'\includegraphics[width=2cm]'\
                   '{mov/wave_frames/frame_%04d.png}\n' %
                       counter)
        elif FORMAT in ("sphinx", "rst", "html"):
           text += ' '\
            '<a href="mov/wave_frames/frame_%04d.png">'\
            '<img src="mov/wave_frames/frame_%04d.png"
               width="300">'\
            '</a> \n' % (counter, counter)
        if fig_counter % no_of_columns == 0: # new row
           if FORMAT in ("latex", "pdflatex"):
    text += ' ' + ' & '.join(latex_columns)
               if counter != end:
                   text += r'\\' + '\n'
               latex_columns = []
            elif FORMAT in ("sphinx", "rst", "html"):
               text += '
'n'
               if counter != end:
                   text += '\n' # begin new row
   # Footer
   if FORMAT in ("latex", "pdflatex"):
       text += r"""\hline
\end{tabular}\end{quote}
   elif FORMAT in ("sphinx", "rst", "html"):
      text += '\n'
```

```
if FORMAT in ("sphinx", "rst"):
    # Wrap raw HTML code
    lines = text.splitlines()
    text = '\n.. raw:: html\n\n'
    for line in lines:
        text += ' ' + line + '\n'
    text += '\n'
    return text
%>
## Example on call in DocOnce source
${generate_native_table(80, 120, 5, 3)}
```

One can easily add support for various other formats, such as pandoc, gwiki, mwiki, etc. The output in format the current format is as above.

Generating the Entire Table in a Native Format as a Post Process. Instead of using Mako as shown above, we can invent our own syntax line for the table with figures and apply a script to the DocOnce output file to replace the special line with proper native code. We use the same <code>generate\_native\_table</code> function as above, but this time in a Python script. In the DocOnce source we place a line

```
@@@FIGTABLE 80 120 5 3
```

which is meant to generate a table with figures where the numbers correspond to arguments in the generative native table function.

Suppose we have run

Terminal> doconce format sphinx mydoc

In the resulting mydoc.rst file we have the magic line @@@FIGTABLE .... Such lines can now be processed in our Python script:

```
# read mydoc.rst into the string filestr
pattern = r'@@@FIGTABLE (.+)'
def replacement(m):
    # m is a MatchObject
    args = m.group(1)
    # Convert args to a tuple of words with right type
    args = [eval(arg) for arg in args.split()]
    text = generate_native_table(*args)

filestr = re.sub(pattern, filestr, replacement)
# write filestr back to manual.rst
```

This code makes use of a function for creating the replacement string for the re.sub command and demands a knowledge of regular expression syntax and the workings or re.sub. A more plain Python version goes like this:

```
# read mydoc.rst into the string filestr
lines = filestr.splitsplines()
for i in range(len(lines)):
    if lines[i].startswith('@@@FIGTABLE'):
```

```
args = lines[i][10:].strip()
    # Convert args to a tuple of words with right type
    args = [eval(arg) for arg in args.split()]
    text = generate_native_table(*args)
    lines[i] = text
filestr = '\n'.join(lines) # make one string from all separate
    lines
# write filestr back to manual.rst
```

#### Recommendation.

This latter way of inserting specialized native text after DocOnce has generated the output file is easier than using Mako and usually also safer.

#### Lesson Learned.

The above examples show different techniques for extending the DocOnce capabilities. The ideas are not restricted to DocOnce: using Mako as a preprocessing step or inserting "magic" lines that you can later substitute by a script is possible in all types of ASCII-based documents, e.g., LATEX, Sphinx, and Markdown.

#### 4.5 Example: Defining a Theorem Environment

DocOnce supports only basic formatting elements (headings, paragraphs, lists, etc.), while LaTeX users are used to fancy environments for, e.g., theorems. A flexible strategy is to typeset theorems using paragraph headings, which will look satisfactorily in all formats, but add comment lines that can be replaced by LaTeX environments via doconce replace. Theorems can be numbered using a variable in Mako. Here is an example on raw DocOnce code:

```
<%
theorem_counter = 4
%>

# begin theorem
label{theorem:fundamental1}
<%
theorem_counter += 1
theorem_fundamental1 = theorem_counter
%>

__Theorem ${theorem_counter}.__
Let $a=1$ and $b=2$. Then $c=3$.
# end theorem
# begin proof
__Proof.__
```

```
Since $c=a+b$, the result follows from straightforward addition.
$\Diamond$|$END$
# end proof
As we see, the proof of Theorem ${theorem_counter} is a modest
achievement.
The .p.tex output file now reads
% begin theorem
label{theorem:fundamental1}
\label{thm:commutation} $$ \operatorname{$5.}$ Let $a=1$ and $b=2$. Then $c=3$.
% end theorem
% begin proof
\paragraph{Proof.}
Since $c=a+b$, the result follows from straightforward addition.
$\Diamond$
% end proof
As we see, the proof of Theorem 5 is a modest
achievement.
```

Note that with Mako variables we can easily create our own counters, and this works in any format. In LATEX we can use both the generated numbers from Mako variables or we can use the labels.

The next step is to replace the % begin ... and % end ... lines with the proper LaTEX expressions in the .p.tex file. Moreover, we need to remove the paragraphs with *Theorem*. The following Bash script does the job:

More heavy editing is better done in a Python script that reads the mydoc.p.tex file and performs string substitutions and regex substitutions as needed.

The resulting mydoc.tex file now becomes

```
\usepackage{theorem}
\newtheorem{theorem}{Theorem} [section]
...
\begin{theorem}
\label{theorem}:fundamental1}

Let $a=1$ and $b=2$. Then $c=3$.
\end{theorem}
% begin proof
```

```
\paragraph{Proof.}
Since $c=a+b$, the result follows from straightforward addition.
$\Diamond$
% end proof
As we see, the proof of Theorem 5 is a modest
achievement.
```

Even better, HTML output looks nice as well.

Note that DocOnce supports fancy environments for verbatim code via the ptex2tex program with all its flexibility for choosing environments. Also doconce ptex2tex has quite some flexibility for typesetting computer code.

# 4.6 Tools for Writing DocOnce Documents

- Emacs (with the modest DocOnce syntax highlighting), Vim, Gedit, etc.
- TeXMaker and Kile can in theory be adapted to handle DocOnce compilation and maybe even DocOnce constructs.

# 4.7 Debugging

Given a problem, extract a small portion of text surrounding the problematic area and debug that small piece of text. DocOnce does a series of transformations of the text. The effect of each of these transformation steps are dumped to a logfile, named \_doconce\_debugging.log, if the to doconce format after the filename is debug. The logfile is inteded for the developers of DocOnce, but may still give some idea of what is wrong. The section "Basic Parsing Ideas" explains how the DocOnce text is transformed into a specific format, and you need to know these steps to make use of the logfile.

#### 5 From DocOnce to Other Formats

Transformation of a DocOnce document mydoc.do.txt to various other formats is done with the script doconce format:

Terminal> doconce format formatname mydoc.do.txt

or just drop the extension:

Terminal > doconce format formatname mydoc

# 5.1 Generating a makefile

Producing HTML, Sphinx, and in particular LaTeX documents from DocOnce sources requires a few commands. Often you want to produce several different formats. The relevant commands should then be placed in a script that acts as a "makefile".

The doconce makefile can be used to automatically generate such a makefile, more precisely a Python script make.py, which carries out the commands explained below. If our DocOnce source is in main\_myproj.do.txt, we run

doconce makefile main\_myproj html pdflatex sphinx

to produce the necessary output for generating HTML, PDFLATEX, and Sphinx. Usually, you need to edit make.py to really fit your needs. Some examples lines are inserted as comments to show various options that can be added to the basic commands. A handy feature of the generated make.py script is that it inserts checks for successful runs of the many doconce commands, and if something goes wrong, the script aborts.

# 5.2 Preprocessing

The preprocess and make programs are used to preprocess the file. The DocOnce program detects whether preprocess and/or make statements are present and runs the corresponding programs, first preprocess and then make.

Variables to preprocess and/or make can be added after the filename with the syntax -DMYVAR, -DMYVAR=val or MYVAR=val.

- The form -DMYVAR defines the variable MYVAR for preprocess (like the same syntax for the C preprocessor - MYVAR is defined, but has not specific value). When running mako, -DMYVAR means that MYVAR has the (Python) value True.
- The expressions -DMYVAR=val and MYVAR=val are equivalent. When running preprocess, MYVAR is defined and has the value val (# #ifdef MYVAR and # #if MYVAR == "val" are both true tests), while for make, MYVAR exists as variable and has the value val (% if MYVAR == "val" is true).

Note that MYVAR=False defines MYVAR in preprocess and any test # #ifdef MYVAR is always true, regardless of the value one has set MYVAR to, so a better test is # #if MYVAR == True. In general, it is recommended to go with preprocess directives if the tests are very simple, as in # #ifdef MYVAR or # #if FORMAT == "latex", otherwise use only make syntax like % if MYVAR or YOURVAR: to incorporate if tests in the preprocessor phases.

Two examples on defining preprocessor variables are

Terminal> doconce format sphinx mydoc -Dextra\_sections -DVAR1=5 Terminal> doconce format sphinx mydoc extra\_sections=True VAR1=5

The variable FORMAT is always defined as the current format when running preprocess or mako. That is, in the above examples, FORMAT is defined as sphinx. Inside the DocOnce document one can then perform format specific actions through tests like #if FORMAT == "sphinx" (for preprocess) or % if FORMAT == "sphinx": (for mako).

The result of running preprocess on a DocOnce file mydoc.do.txt is available in a file tmp\_preprocess\_mydoc.do.txt. Similarly, the result of running mako is available in tmp\_mako\_mydoc.do.txt. By examining these files one can see exactly what the preprocessors have done.

#### 5.3 Removal of inline comments

The command-line arguments --no\_preprocess and --no\_make turn off running preprocess and make, respectively.

Inline comments in the text are removed from the output by

Terminal> doconce format latex mydoc --skip\_inline\_comments

One can also remove all such comments from the original DocOnce file by running:

Terminal> doconce remove\_inline\_comments mydoc

This action is convenient when a DocOnce document reaches its final form and comments by different authors should be removed.

#### 5.4 Notes

DocOnce does not have a tag for longer notes, because implementation of a "notes feature" is so easy using the preprocess or make programs. Just introduce some variable, say NOTES, that you define through -DNOTES (or not) when running doconce format .... Inside the document you place your notes between # #ifdef NOTES and # #endif preprocess tags. Alternatively you use % if NOTES: and % endif that make will recognize. In the same way you may encapsulate unfinished material, extra material to be removed for readers but still nice to archive as part of the document for future revisions.

#### 5.5 Demo of different formats

A simple scientific report is available in a lot of different formats. How to create the different formats is explained in more depth in the coming sections.

#### **5.6 HTML**

**Basics.** Making an HTML version of a DocOnce file mydoc.do.txt is performed by

Terminal> doconce format html mydoc

The resulting file mydoc.html can be loaded into any web browser for viewing.

**Typesetting of Code.** If the Pygments package (including the pygmentize program) is installed, code blocks are typeset with aid of this package. The command-line argument --no\_pygments\_html turns off the use of Pygments and makes code blocks appear with plain (pre) HTML tags. The option --pygments\_html\_linenos turns on line numbers in Pygments-formatted code blocks. A specific Pygments style is set by --pygments\_html\_style=style, where style can be default, emacs, perldoc, and other valid names for Pygments styles.

Handling of Movies. MP4, WebM, and Ogg movies are typeset with the HTML5 video tag and the HTML code tries to load as many versions among MP4, WebM, and Ogg as exist (and the files are loaded in the mentioned order). If just the specified file is to be loaded, use the --no\_mp4\_webm\_ogg\_alternatives command-line option. Other movie formats, e.g., .flv, .mpeg and .avi, are embedded via the older embed tag.

**HTML Styles.** The HTML style can be defined either in the header of the HTML file, using a named built-in style; in an external CSS file; or in a template file.

An external CSS file filename used by setting the command-line argument --css=filename. There available built-in styles are specified as --html\_style=name, where name can be

- solarized: the famous solarized style (yellowish),
- blueish: a simple style with blue headings (default),
- blueish2: a variant of bluish,
- bloodish: as bluish, but dark read as color,
- bootstrap\* or bootswatch\* in a lot of variants, see doconce format --help for a list of all styles.

There is a comprehensive demonstration of almost all available styles!

Using --css=filename where filename is a non-existing file makes Do-

cOnce write the built-in style to that file. Otherwise the HTML links to the CSS stylesheet in filename. Several stylesheets can be specified: --ccs=file1.css,file2.css,file3.css.

**HTML templates.** Templates are HTML files with "slots" %(main)s for the main body of text, %(title)s for the title, and %(date)s for the date. DocOnce comes with a few templates. The usage of templates is described in a "separate document": "". That document describes how you your DocOncegenerated HTML file can have any specified layout.

The HTML file can be embedded in a template with your own tailored design, see a tutorial on this topic. The template file must contain valid HTML code and can have three "slots": "(title)s for a title, "(date)s for a date, and

%(main)s for the main body of text. The latter is the DocOnce document translated to HTML. The title becomes the first heading in the DocOnce document, or the title (but a title is not recommended when using templates). The date is extracted from the DATE: line. With the template feature one can easily embed the text in the look and feel of a website. DocOnce comes with two templates in bundled/html\_styles. Just copy the directory containing the template and the CSS and JavaScript files to your document directory, edit the template as needed (also check that paths to the css and js subdirectories are correct according to how you store the template files), and run

```
Terminal> doconce format html mydoc --html_template=mytemplate.html
```

The template in style\_vagrant also needs an extra option --html\_style=vagrant.

The HTML File Collection. There are usually a range of files needed for an HTML document arising from a DocOnce source. The needed files are listed in .basename\_html\_file\_collection, where basename is the filestem of the DocOnce file (i.e., the DocOnce source is in basename.do.txt).

**Filenames.** An HTML version of a DocOnce document is often made in different styles, calling for a need to rename the HTML output file. This is conveniently done by the --html\_output=basename option, where basename is the filestem of the associated HTML files. The .basename\_html\_file\_collection file lists all the needed files for the HTML document. Here is an example on making three versions of the HTML document: mydoc\_bloodish.html, mydoc\_solarized, and mydoc\_vagrant.

# 5.7 Blog Posts

DocOnce can be used for writing blog posts provided the blog site accepts raw HTML code. Google's Blogger service (blogger.com or blogname.blogspot.com) is particularly well suited since it also allows extensive LATEX mathematics via MathJax.

- 1. Write the text of the blog post as a DocOnce document without any title, author, and date.
- 2. Generate HTML as described above.

3. Copy the text and paste it into the text area in the blog post (just delete the HTML code that initially pops up in the text area). Make sure the input format is HTML.

See a simple blog example and a scientific report for demonstrations of blog posts at blogspot.no.

#### Warning.

The comment field after the blog post does not recognize MathJax (LTEX) mathematics or code with indentation. However, using a MathJax bookmarklet, e.g., at http://checkmyworking.com/misc/mathjax-bookmarklet/one can get the mathematics properly rendered. The comment fields are not suitable for computer code, though, as HTML tags are not allowed.

#### Notice.

Figure files must be uploaded to some web site and the local filenames name must be replaced by the relevant URL. This is usually done by using the --figure\_prefix=http://project.github.io/... option to give some URL as prefix to all figure names (a similar --movie\_prefix= option exists as well).

Changing figure names in a blog post can also be done "manually" by some editing code in the script that compiles the DocOnce document to HTML format:

Blog posts at Google can also be published automatically through email. A Python program can send the contents of the HTML file to the blog site's email address using the packages smtplib and email.

WordPress (wordpress.com) allows raw HTML code in blogs, but has very limited LaTeX support, basically only formulas. The --wordpress option to doconce modifies the HTML code such that all equations are typeset in a way that is acceptable to WordPress. Look at a simple doconce example and a scientific report to see blog posts with mathematics and code on WordPress.

Speaking of WordPress, the related project http://pressbooks.com can take raw HTML code (from DocOnce, for instance) and produce very nicelooking books. There is no support for mathematics in the text, though.

#### 5.8 Pandoc and Markdown

Output in Pandoc's extended Markdown format results from

```
Terminal> doconce format pandoc mydoc
```

The name of the output file is mydoc.md. From this format one can go to numerous other formats:

```
Terminal> pandoc -R -t mediawiki -o mydoc.mwk --toc mydoc.md
```

Pandoc supports latex, html, odt (OpenOffice), docx (Microsoft Word), rtf, texinfo, to mention some. The -R option makes Pandoc pass raw HTML or Latex to the output format instead of ignoring it, while the --toc option generates a table of contents. See the Pandoc documentation for the many features of the pandoc program.

Markdown to HTML conversion. The HTML output from pandoc needs adjustments to provide full support for MathJax LaTEX mathematics, and for this purpose one should use doconce md2html:

```
Terminal> doconce format pandoc mydoc Terminal> doconce m2html mydoc
```

The result mydoc.html can be viewed in a browser.

Using Pandoc to go from LaTeX to MS Word or HTML. Pandoc is useful to go from LaTeX mathematics to, e.g., HTML or MS Word. There are two ways (experiment to find the best one for your document): doconce format pandoc and then translating using doconce md2latex (which runs pandoc), or doconce format latex, and then going from LaTeX to the desired format using pandoc. Here is an example on the latter strategy:

```
Terminal> doconce format latex mydoc
Terminal> doconce ptex2tex mydoc
Terminal> doconce replace '\Verb!' '\verb!' mydoc.tex
Terminal> pandoc -f latex -t docx -o mydoc.docx mydoc.tex
```

When we go through pandoc, only single equations, align, or align\* environments are well understood for output to HTML.

Note that DocOnce applies the Verb macro from the fancyvrb package while pandoc only supports the standard verb construction for inline verbatim text. Moreover, quite some additional doconce replace and doconce subst edits might be needed on the .mkd or .tex files to successfully have mathematics that is well translated to MS Word. Also when going to reStructuredText using Pandoc, it can be advantageous to go via LATEX.

**Strict Markdown.** The option --strict\_markdown\_output generates plain or strict Markdown without the many extension that Pandoc accepts in Markdown syntax.

**GitHub-flavored Markdown.** Adding the command-line option <code>github-md</code> turns on the GutHub-flavored Markdown dialect, which is used for the issue tracker on GitHub. A special feature is the support of task lists: unnumbered lists with [x] (task done) or  $[\ ]$  (task not done). (Tables get typeset directly as HTML and the syntax for code highlighting is different from Pandoc extended Markdown.) Below is a typical response in a GitHub issue tracker where one first quotes the issue and then provides an answer:

```
There is a problem with the 'f(x)' function
!bc pycod def f(x):
  return 1 + x
This function should be quadratic.
!equote
OK, this is fixed:
!bc pycod
def f(x, a=1, b=1, c=1):
    return a*x**2 + b*x + c
===== Updated task list =====
   * [x] Offer an 'f(x)' function
   * [ ] Extension to cubic functions
   * [x] Allowing general coefficient in the quadratic function
=== Remaining functionality ===
  function | purpose
                      | state |
              ---1-----
   'g(x)' | Compute the Gaussian function. | Formula ready. |
'h(x)' | Heaviside function. | Formula ready. |
'I(x)' | Indicator function. | Nothing done yet. |
```

!bauote

==== Problems with a function =====

Say this text is stored in a file mycomments.do.txt. Running

Terminal> doconce format pandoc mycomments --github\_md

produces the Markdown file mycomments.md, which can be pasted into the Write field of the GitHub issue tracker. Turning on Preview shows the typesetting of the quote, compute code, inline verbatim, headings, the task list, and the table.

**MultiMarkdown.** The option --multimarkdown\_output generates the Multi-Markdown version of Markdown (as opposed to Pandoc-extended Markdown (default), strict Markdown, or GitHub-flavored Markdown).

Strapdown rendering of Markdown text. Strapdown is a tool that can render Markdown text nicely in a web browser by just inserting an HTML header and footer in the Markdown file and load the file into a browser. The option—strapdown outputs the relevant header and footer. The output file must be renamed such that it gets the extension .html:

The --bootstrap\_bootwatch\_theme=theme option is used to choose a Bootswatch theme whose names are found on the Strapdown page.

# 5.9 LAT<sub>E</sub>X

#### Notice.

XeLaTeX and PDFLATEX are used very much in the same way as standard LATEX. The minor differences are described in separate sections of the documentation of the DocOnce to LATEX translation.

Making a LaTeX file mydoc.tex from mydoc.do.txt is done in two steps: 1) compile the DocOnce source to the ptex2tex format, and 2) compile the ptex2tex format to standard LaTeX. The ptex2tex format can be viewed as an extended LaTeX. For DocOnce users, the ptex2tex format essentially means that the file consists of

- 1. if-else statements for the preprocess processor such that LATEX constructions can be activated or deactivated, and
- all code environments can be typeset according to a .ptex2tex.cfg configuration file.

Point 2 is only of interest if you aim to use a special computer code formatting that requires you to use a configuration file and the ptex2tex program.

The reason for generating ptex2tex and not standard LTEX directly from DocOnce was that the ptex2tex format shows a range of possible LTEX constructions for controlling the layout. It can be instructive for LTEX users to look at this code before choosing specific parts for some desired layout. Experts may also want to edit this code (which should be automated by a script such that the edits can be repeated when the DocOnce source is modified, see Step 2b below). (Direct control of the LTEX layout in the doconce format program would

not spit out alternative LATEX constructs as is now done through the ptex2tex step.)

Going from ptex2tex format to standard LaTeX format is enabled by either the ptex2tex program or DocOnce's (simplified) version of it: doconce ptex2tex. Details are given below.

#### Information on typesetting of inline verbatim.

The ptex2tex and the doconce ptex2tex programs take inline verbatim code, typeset with backticks in DocOnce, and translate this to

\Verb!text!

Thereafter, if text does not contain illegal characters for the \texttt command, the latter is used instead since then LATEX can insert linebreaks in the inline verbatim text and hence avoid overfull hboxes.

#### **Step 1.** Filter the doconce text to the ptex2tex "pre-LATEX form" mydoc.p.tex:

Terminal > doconce format latex mydoc

LATEX-specific commands ("newcommands") in math formulas and similar can be placed in files newcommands.tex, newcommands\_keep.tex, or newcommands\_replace.tex (see Section 2.20). If these files are present, they are included in the LATEX document so that your commands are defined.

An option --device=paper makes some adjustments for documents aimed at being printed. For example, links to web resources are associated with a footnote listing the complete web address (URL). (Very long URLs in footnotes can be shortened using services such as http://goo.gl/, http://tinyurl.com/, and https://bitly.com/.) The default, --device=screen, creates a PDF file for reading on a screen where links are just clickable.

There are many additional options (run doconce format --help and look for options starting with --latex to get a more verbose description):

- --latex\_font=helvetica,palatino
- --latex\_papersize=a4,a6
- --latex\_bibstyle=plain
- --latex\_title\_layout=titlepage,std,beamer,doconce\_heading
- --latex\_style=std,Springer\_lncse,Springer\_llncs,Springer\_T2,...
- --latex\_list\_of\_exercises=loe,toc,none
- --latex\_fancy\_header

- --latex\_section\_headings=std,blue,strongblue,gray,gray-wide
- --latex\_colored\_table\_rows=blue,gray,no
- --latex todonotes
- --latex\_double\_spacing
- --latex preamble=filename
- --latex\_admon=mdfbox,graybox2,grayicon,yellowicon,paragraph,colors1,colors2
- --latex\_admon\_color=0.34,0.02,0.8
- --latex\_admon\_envir\_map=2
- --latex\_exercise\_numbering=absolute,chapter
- --latex\_movie=media9,href,multimedia,movie15
- --latex\_movie\_controls=on
- --latex\_external\_movie\_viewer (for movie15 package)
- --xelatex

**Step 2.** Run ptex2tex (if you have installed the Python ptex2tex package) to make a standard LATEX file,

Terminal> ptex2tex mydoc

In case you do not have ptex2tex, you may run the (simplified) version that comes with DocOnce:

Terminal> doconce ptex2tex mydoc

The ptex2tex command can set two preprocessor variables:

- PREAMBLE to turn the LATEX preamble on or off (i.e., complete document or document to be included elsewhere - and note that the preamble is only included if the document has a title, author, and date)
- MINTED for inclusion of the minted package for typesetting of code with the Pygments tool (which requires latex or pdflatex to be run with the -shell-escape option); not used for doconce ptex2tex only in the ptex2tex program

If you are not satisfied with the generated DocOnce preamble, you can provide your own preamble by adding the command-line option --latex\_preamble=myfile. In case myfile contains a documentclass definition, DocOnce assumes that the file contains the *complete* preamble you want (not that all the packages listed in the default preamble are required and must be present in myfile). Otherwise, myfile is assumed to contain *additional* LaTeX code to be added to the DocOnce default preamble.

The ptex2tex tool makes it possible to easily switch between many different fancy formattings of computer code in LaTeX documents. After any !bc command in the DocOnce source you can insert verbatim block styles as defined in your .ptex2tex.cfg file, e.g., !bc sys for a terminal session, where sys is set to a certain environment in .ptex2tex.cfg (e.g., CodeTerminal). There are about 40 styles to choose from, and you can easily add new ones.

The doconce ptex2tex allows specifications of code environments as well. Here is an example:

Note that @ must be used to separate the begin and end LTEX commands, unless only the environment name is given (such as minted above, which implies \begin{minted}{fortran} and \end{minted} as begin and end for blocks inside !bc fpro and !ec). Specifying envir=ans:nt means that all other environments are typeset with the anslistings.sty package, e.g., !bc cppcod will then result in \begin{c++}. A predefined shortcut as in shcod=Verbatim-0.85 results in denser vertical spacing (baselinestretch 0.85 in LTEX terminology), and shcod=Verbatim-indent implies indentation of the verbatim text. Alternatively, one can provide all desired parameters \begin{Verbatim} instruction using the syntax illustrated for the sys environments above.

If no environments like sys, fpro, or the common envir are defined on the command line, the plain \begin{Verbatim} and \end{Verbatim} instructions are used.

**Step 2b (optional).** Edit the mydoc.tex file to your needs. For example, you may want to substitute section by section\* to avoid numbering of sections, you may want to insert linebreaks (and perhaps space) in the title, etc. This can be automatically edited with the aid of the doconce replace and doconce subst commands. The former works with substituting text directly, while the latter performs substitutions using regular expressions. You will use doconce replace to edit section\* {

```
Terminal> doconce replace 'section{' 'section*{' mydoc.tex
```

For fixing the line break of a title, you may pick a word in the title, say "Using", and insert a break after than word. With doconce subst this is easy employing regular expressions with a group before "Using" and a group after:

```
Terminal> doconce subst 'title\{(.+)Using (.+)\}' \
    'title\\g<1> \\\\ [1.5mm] Using \g<2>' mydoc.tex
```

A lot of tailored fixes to the LaTeX document can be done by an appropriate set of text replacements and regular expression substitutions. You are anyway encourged to make a script for generating PDF from the LaTeX file so the doconce subst or doconce replace commands can be put inside the script.

#### Step 3. Compile mydoc.tex and create the PDF file:

```
Terminal> latex mydoc
Terminal> latex mydoc
Terminal> makeindex mydoc  # if index
Terminal> bibitem mydoc  # if bibliography
Terminal> latex mydoc
Terminal> dvipdf mydoc
```

See the next two sections for compilation with XeLaTeX or PDFLATEX.

If one wishes to use the minted LaTeX package for typesetting code blocks (Minted\_Python, Minted\_Cpp, etc., in ptex2tex specified through the \*pro and \*cod variables in .ptex2tex.cfg or \$HOME/.ptex2tex.cfg), the minted LaTeX package is needed. This package is automatically included by doconce ptex2tex if the minted style is used, while you have to include the -DMINTED preprocessor option when running the ptex2tex program:

```
Terminal> ptex2tex -DMINTED mydoc
```

If the minted style is used, latex (or pdflatex or xelatex) *must* be run with the -shell-escape option:

```
Terminal> latex -shell-escape mydoc
Terminal> latex -shell-escape mydoc
Terminal> makeindex mydoc  # if index
Terminal> bibitem mydoc  # if bibliography
Terminal> latex -shell-escape mydoc
Terminal> dvipdf mydoc
```

#### 5.10 PDFLaTeX

Running pdflatex instead of latex follows almost the same steps, but the start is

```
Terminal> doconce format latex mydoc
```

Then ptex2tex is run as explained above, and finally

```
Terminal> pdflatex -shell-escape mydoc
Terminal> makeindex mydoc  # if index
Terminal> bibitem mydoc  # if bibliography
Terminal> pdflatex -shell-escape mydoc
```

#### 5.11 XeLaTeX

XeLaTeX is an alternative to PDFLTEX and is run in almost the same way, except for the --xelatex flag to doconce format:

```
Terminal> doconce format pdflatex mydoc --xelatex Terminal> doconce ptex2tex mydoc Terminal> xelatex mydoc
```

## 5.12 From PDF to e-book formats

PDF (as generated from LaTeX above) can be read on most devices today. However, for Kindle and other devices specialized for e-books you need to convert to their format. The Calibre program can produce epub, mobi, and other e-book formats from PDF, see a description.

## 5.13 Plain ASCII Text

We can go from DocOnce "back to" plain untagged text suitable for viewing in terminal windows, inclusion in email text, or for insertion in computer source code:

```
Terminal> doconce format plain mydoc.do.txt # results in mydoc.txt
```

#### 5.14 reStructuredText

Going from DocOnce to reStructuredText gives a lot of possibilities to go to other formats. First we filter the DocOnce text to a reStructuredText file mydoc.rst:

```
Terminal> doconce format rst mydoc.do.txt
```

We may now produce various other formats:

```
Terminal> rst2html.py mydoc.rst > mydoc.html # html
Terminal> rst2latex.py mydoc.rst > mydoc.tex # latex
Terminal> rst2xml.py mydoc.rst > mydoc.xml # XML
Terminal> rst2odt.py mydoc.rst > mydoc.odt # OpenOffice
```

The OpenOffice file mydoc.odt can be loaded into OpenOffice and saved in, among other things, the RTF format or the Microsoft Word format. However, it is more convenient to use the program unovonv to convert between the many formats OpenOffice supports on the command line. Run

```
Terminal> unoconv --show
```

to see all the formats that are supported. For example, the following commands take mydoc.odt to Microsoft Office Open XML format, classic MS Word format, and PDF:

```
Terminal> unoconv -f ooxml mydoc.odt
Terminal> unoconv -f doc mydoc.odt
Terminal> unoconv -f pdf mydoc.odt
```

Remark about Mathematical Typesetting. At the time of this writing, there is no easy way to go from DocOnce and LaTeX mathematics to reST and further to OpenOffice and the "MS Word world". Mathematics is only fully supported by latex as output and to a wide extent also supported by the sphinx output format. Some links for going from LaTeX to Word are listed below.

- http://ubuntuforums.org/showthread.php?t=1033441
- http://tug.org/utilities/texconv/textopc.html
- http://nileshbansal.blogspot.com/2007/12/latex-to-openofficeword.
   html

## 5.15 Sphinx

Sphinx documents demand quite some steps in their creation. We have automated most of the steps through the doconce sphinx\_dir command:

The keywords author, title, and version are used in the headings of the Sphinx document. By default, version is 1.0 and the script will try to deduce authors and title from the doconce file mydoc.do.txt The default value of dirname is sphinx-rootdir. The theme keyword is used to set the theme for design of HTML output from Sphinx (the default theme is 'default').

One often just runs the simple command

Terminal > doconce sphinx\_dir mydoc

which creates the Sphinx directory sphinx-rootdir with relevant files.

The doconce sphinx\_dir command generates a script automake\_sphinx.py for compiling the Sphinx document into an HTML document. Run

```
Terminal> python automake_sphinx.py
```

As the output also tells, you can see the Sphinx HTML version of the document by running

```
Terminal> google-chrome sphinx-rootdir/_build/html/index.html
```

or loading the index.html file manually into your favorite web browser.

If you cycle through editing the DocOnce file and watching the HTML output, you should observe that automake\_sphinx.py does not recompile the DocOnce file if the Sphinx .rst version already exists. In each edit-and-watch cycle do

```
Terminal> rm mydoc.rst; python automake_sphinx.py
```

#### Tip.

If you are new to Sphinx and end up producing quite some Sphinx documents, you are encouraged to read the Sphinx documentation and study the automake sphinx.py file. Maybe you want to do things differently.

The following paragraphs describes the many possibilities for steering the Sphinx output.

**Links.** The automake\_sphinx.py script copies directories named fig\* over to the Sphinx directory so that figures are accessible in the Sphinx compilation. It also examines MOVIE: and FIGURE: commands in the DocOnce file to find other image files and copies these too. I strongly recommend to put files to which there are local links (not http: or file: URLs) in a directory named \_static. The automake\_sphinx.py copies \_static\* to the Sphinx directory, which guarantees that the links to the local files will work in the Sphinx document.

There is a utility doconce sphinxfix\_localURLs for checking links to local files and moving the files to \_static and changing the links accordingly. For example, a link to dir1/dir2/myfile.txt is changed to \_static/myfile.txt and myfile.txt is copied to \_static. However, I recommend instead that you manually copy files to \_static when you want to link to them, or let your script which compiles the DocOnce document do it automatically.

**Themes.** DocOnce comes with a rich collection of HTML themes for Sphinx documents, much larger than what is found in the standard Sphinx distribution. Additional themes include agni, basicstrap, bootstrap, cloud, fenics, fenics\_minimal, flask, haiku, impressjs, jal, pylons, redcloud, scipy\_lectures, slim-agogo, and vlinux-theme.

All the themes are packed out in the Sphinx directory, and the doconce sphinx\_dir insert lots of extra code in the conf.py file to enable easy specification and customization of themes. For example, modules are loaded for the additional themes that come with DocOnce, code is inserted to allow customization of the look and feel of themes, etc. The conf.py file is a good starting point for fine-tuning your favorite team, and your own conf.py file can later be supplied and used when running doconce sphinx\_dir: simply add the command-line option conf.py=conf.py.

A script make-themes.sh can make HTML documents with one or more themes. For example, to realize the themes fenics, pyramid, and pylon one writes

Terminal> ./make-themes.sh fenics pyramid pylon

The resulting directories with HTML documents are \_build/html\_fenics and \_build/html\_pyramid, respectively. Without arguments, make-themes.sh makes

all available themes (!). With make-themes.sh it is easy to check out various themes to find the one that is most attractive for your document.

You may supply your own theme and avoid copying all the themes that come with DocOnce into the Sphinx directory. Just specify theme\_dir=path on the command line, where path is the relative path to the directory containing the Sphinx theme. You must also specify a configure file by conf.py=path, where path is the relative path to your conf.py file.

**Example.** Say you like the scipy\_lectures theme, but you want a table of contents to appear to the right, much in the same style as in the default theme (where the table of contents is to the left). You can then run doconce sphinx\_dir, invoke a text editor with the conf.py file, find the line html\_theme == 'scipy\_lectures', edit the following nosidebar to false and rightsidebar to true. Alternatively, you may write a little script using doconce replace to replace a portion of text in conf.py by a new one:

```
doconce replace "elif html_theme == 'scipy_lectures':
    html_theme_options = {
        'nosidebar': 'true',
        'rightsidebar': 'false'
        'sidebarbgcolor': '#f2f2f2',
        'sidebartextcolor': '#20435c',
        'sidebarlinkcolor': '#20435c',
        'footerbgcolor': '#000000',
        'relbarbgcolor': '#000000'.
    }" "elif html_theme == 'scipy_lectures':
    html_theme_options = {
        'nosidebar': 'false',
        'rightsidebar': 'true'
        'sidebarbgcolor': '#f2f2f2',
        'sidebartextcolor': '#20435c',
        'sidebarlinkcolor': '#20435c',
        'footerbgcolor': '#000000',
        'relbarbgcolor': '#000000',
    }" conf.py
```

Obviously, we could also have changed colors in the edit above. The final alternative is to save the edited conf.py file somewhere and reuse it the next time doconce  $sphinx_dir$  is run

The manual Sphinx procedure. If it is not desirable to use the autogenerated scripts explained above, here is the complete manual procedure of generating a Sphinx document from a file mydoc.do.txt.

#### **Step 1.** Translate DocOnce into the Sphinx format:

Terminal> doconce format sphinx mydoc

**Step 2.** Create a Sphinx root directory either manually or by using the interactive sphinx-quickstart program. Here is a scripted version of the steps with the latter:

```
mkdir sphinx-rootdir
sphinx-quickstart <<EOF
sphinx-rootdir
Name of My Sphinx Document
Author
version
version
.rst
index
У
n
n
n
n
У
n
у
y
EOF
```

The autogenerated <code>conf.py</code> file may need some edits if you want to specific layout (Sphinx themes) of HTML pages. The <code>doconce sphinx\_dir</code> generator makes an extended <code>conv.py</code> file where, among other things, several useful Sphinx extensions are included.

#### **Step 3.** Copy the mydoc.rst file to the Sphinx root directory:

Terminal> cp mydoc.rst sphinx-rootdir

If you have figures in your document, the relative paths to those will be invalid when you work with mydoc.rst in the sphinx-rootdir directory. Either edit mydoc.rst so that figure file paths are correct, or simply copy your figure directories to sphinx-rootdir. Links to local files in mydoc.rst must be modified to links to files in the \_static directory, see comment above.

**Step 4.** Edit the generated index.rst file so that mydoc.rst is included, i.e., add mydoc to the toctree section so that it becomes

```
.. toctree::
:maxdepth: 2
mydoc
```

(The spaces before mydoc are important!)

#### **Step 5.** Generate, for instance, an HTML version of the Sphinx source:

```
make clean  # remove old versions
make html
```

Sphinx can generate a range of different formats: standalone HTML, HTML in separate directories with index.html files, a large single HTML file, JSON files, various help files (the qthelp, HTML, and Devhelp projects), epub, LTEX, PDF (via LTEX), pure text, man pages, and Texinfo files.

## Step 6. View the result:

```
Terminal> firefox _build/html/index.html
```

Note that verbatim code blocks can be typeset in a variety of ways depending the argument that follows !bc: cod gives Python (code-block:: python in Sphinx syntax) and cppcod gives C++, but all such arguments can be customized both for Sphinx and LaTeX output.

#### 5.16 Wiki Formats

There are many different wiki formats, but DocOnce only supports three: Google-code wiki, MediaWiki, and Creole Wiki. These formats are called gwiki, mwiki, and cwiki, respectively. Transformation from DocOnce to these formats is done by

```
Terminal> doconce format gwiki mydoc.do.txt
Terminal> doconce format mwiki mydoc.do.txt
Terminal> doconce format cwiki mydoc.do.txt
```

The produced MediaWiki can be tested in the sandbox of wikibooks.org. The format works well with Wikipedia, Wikibooks, and ShoutWiki, but not always well elsewhere (see this example).

Large MediaWiki documents can be made with the Book creator. From the MediaWiki format one can go to other formats with aid of mwlib. This means that one can easily use DocOnce to write Wikibooks and publish these in PDF and MediaWiki format, while at the same time, the book can also be published as a standard LATEX book, a Sphinx web document, or a collection of HTML files.

The Googlecode wiki document, mydoc.gwiki, is most conveniently stored in a directory which is a clone of the wiki part of the Googlecode project. This is far easier than copying and pasting the entire text into the wiki editor in a web browser.

When the DocOnce file contains figures, each figure filename must in the .gwiki file be replaced by a URL where the figure is available. There are instructions in the file for doing this. Usually, one performs this substitution automatically (see next section).

## 5.17 Google Docs

Google Docs are normally made online in the interactive editor. However, you may upload a DocOnce document to Google Docs. This requires transforming the DocOnce document to one of the accepted formats for Google Docs:

- OpenOffice: doconce format rst and then run rst2odt (or rst2odt.py).
   Upload the .odt file, click Open... in Google Drive and choose Google Docs as viewer.
- MS Word: doconce format pandoc and then run pandoc to produce a .docx file that can be uploaded to Google Drive and opened in Google Docs.
- RTF: doconce format pandoc and then run pandoc to produce a .rtf file that can be uploaded to Google Drive and opened. Another possibility is to run doconce format latex and then latex2rtf (the support of mathematics has gotten worse).
- Plain text: doconce format plain. Upload the .txt file to Google Drive and open in Google Docs.
- HTML: doconce format html. Upload the .html file and open in Google Docs. Complicated HTML files can be misinterpreted by Google Docs.

This is not yet much tested. It remains to see how code becomes in Google Docs. Support for mathematics is probably impossible until Google Docs can import LaTeX files, but LaTeX mathematics can be embedded in Google Docs and the googledoc2latex script can convert a Google document to LaTeX.

## 5.18 Tweaking the DocOnce Output

Occasionally, one would like to tweak the output in a certain format from DocOnce. One example is figure filenames when transforming DocOnce to re-StructuredText. Since DocOnce does not know if the .rst file is going to be filtered to LaTeX or HTML, it cannot know if .eps or .png is the most appropriate image filename. The solution is to use a text substitution command or code with, e.g., sed, perl, python, or scitools subst, to automatically edit the output file from DocOnce. It is then wise to run DocOnce and the editing commands from a script to automate all steps in going from DocOnce to the final format(s). The make.sh files in docs/manual and docs/tutorial constitute comprehensive examples on how such scripts can be made.

# 6 Options for the doconce commands

## 6.1 doconce format command-line options

The transformation of a DocOnce source to various format is done with the doconce format command, which has a lot of command-line options. These are printed out by doconce format --help. The output is listed here for convenience.

```
Terminal> doconce format --help
doconce format X doconcefile
where X can be any of the formats
html, latex, pdflatex, rst, sphinx, plain, gwiki, mwiki, cwiki,
pandoc, epytext.
--help
Print all options to the doconce program.
--debug
Write a debugging file _doconce_debugging.log with lots
of intermediate results
--no_abort
Do not abort the execution if syntax errors are found.
--skip_inline_comments
Remove all inline comments of the form [ID: comment].
--encoding=...
Specify encoding (e.g., latin1 or utf-8).
--no_mako
Do not run the Mako preprocessor program.
--no_preprocess
Do not run the Preprocess preprocessor program.
--mako_strict_undefined
Make Mako report on undefined variables.
--no_header_footer
```

Do not include header and footer in (LaTeX and HTML) documents.

#### --runestone

Make a RunestoneInteractive version of a Sphinx document.

```
--max_bc_linelength=...
```

Strip lines in !bc environments that are longer than specified (to prevent too long lines). Default: None (no length restriction).

#### --keep\_pygments\_html\_bg

Do not allow change of background in code blocks in HTML.

```
--minted_latex_style=...
```

Specify the minted style to be used for typesetting code in LaTeX.

```
--pygments_html_style=...
```

Specify the minted/pygments style to be used for typesetting code in HTML. Default: default (other values: monokai, manni, rrt, perldoc, borland, colorful, murphy, trac, tango, fruity, autumn, emacs, vim, pastie, friendly, native). none, no, off: turn off pygments to typeset computer code in HTML, use plain tags

#### --pygments\_html\_linenos

Turn on line numbers in pygmentized computer code in HTML. (In LaTeX line numbers can be added via doconce subst or doconce replace such that the verbatim environments get the linenos=true parameter.)

#### --html\_output=...

Alternative basename of files associated with the HTML format.

#### --html\_style=...

#### --html\_code\_style=...

off, inherit, transparent: enable normal inline verbatim font

```
surroundnings (e.g., to avoid the red Boostrap color).
Default: on.
--html_pre_style=...
off, inherit, transparent: let code blocks inside  tags have
foreground and background color inherited from the surroundnings.
Default: on.
--html_template=...
Specify an HTML template with header/footer in which the doconce
document is embedded.
--html_toc_depth=...
No of levels in the table of contents in HTML output for
Bootstrap-based styles. Default: 2.
--html_toc_indent=...
No of spaces for indentation of subsections in the table of
contents in HTML output. Default: 3 (0 gives toc as nested list
in Bootstrap-based styles).
--html_body_font=...
Specify HTML font for text body. =? lists available fonts.
--html_heading_font=...
Specify HTML font for headings. =? lists available fonts.
--html_video_autoplay=...
True for autoplay when HTML is loaded, otherwise False (default).
--html_admon=...
Type of admonition and color:
colors, gray, yellow, apricot, lyx, paragraph.
For html_style=bootstrap*,bootswatch*,
the two legal values are boostrap_panel, bootstrap_alert.
--html_admon_shadow
Add a shadow effect to HTML admon boxes (gray, yellow, apricot).
--html_admon_bg_color=...
Background color of admon in HTML.
```

where foreground and background color is inherited from the

```
--html_admon_bd_color=...
Boundary color of admon in HTML.
--css=...
Specify a .css style file for HTML output.
If the file does not exist, the default or specified style
(--html_style=) is written to it.
--html_box_shadow
Add a shadow effect in HTML box environments.
--html_slide_theme=...
Option for doconce slides_html: specify a theme for the
present slide type.
Examples from reveal.js: darkgray, beige, solarized, default;
from deck.js: sandstone.default, swiss, web-2.0.
(The HTML file contains a comment with the various stylesheets
for the various available themes.)
--html_footer_logo=...
Specify a filename or a style name for a logo in the slide footer.
--beamer_slide_theme=...
Specify a theme for beamer slides. Examples:
--html_exercise_icon=...
Specify a question icon in bundled/html_images for being
inserted to the right in exercises - "default" and "none" are allowed
("none" if no option).
--html_exercise_icon_width=...
Width of the icon image specified as --html_exercise_icon.
--html_links_in_new_window
Open HTML links in a new window.
--html_quiz_button_text=...
Text on buttons for collapsing/expanding answers and
explanations
in quizzes (with bootstrap styles).
Default: Empty (just pencil glyphion).
--html_bootstrap_navbar=...
```

```
Turns the Bootstrap navigation bar on/off. Default: on.
--html_bootstrap_jumbotron=...
Turns the Bootstrap jumbotron intro on/off and governs the size of the heading. Default: on. Other values: h2, off
(h2 gives h2 heading instead of h1, off gives no jumbotron).
--device=...
Set device to paper, screen, or other (paper impacts LaTeX output).
--latex_style=...
LaTeX style package used for the document.
std: standard LaTeX article or book style,
Springer_lncse: Springer's Lecture Notes in Computational Science and
Engineering (LNCSE) style,
Springer_llncs: Springer's Lecture Notes in Computer Science style,
Springer_T2: Springer's T2 book style,
Springer_collection: Springer's style for chapters in LNCSE proceedings, Korma_Script: Korma Script style, siamltex: SIAM's standard LaTeX style for papers,
siamltexmm: SIAM's extended (blue) multimedia style for papers.
--latex_font=...
LaTeX font choice: helvetica, palatino, std (Computer Modern, default).
--latex_bibstyle=...
LaTeX bibliography style. Default: plain.
--section_numbering=...
Turn section numbering on/off. Default: off for all formats except latex and pdflatex (on for the
--latex_table_align=...
Default: quote. Other values: left, center.
--latex_title_layout=...
Layout of the title, authors, and date:
std: traditional LaTeX layout,
titlepage: separate page
doconce_heading (default): authors with "footnotes" for institutions,
beamer: layout for beamer slides.
--latex_papersize=...
Geometry of page size: a6, a4, std (default).
--latex_list_of_exercises=...
```

```
LaTeX typesetting of list of exercises:
loe: special, separate list of exercises,
toc: exercises included as part of the table of contents, none (default): no list of exercises.
--latex_movie=...
Specify package for handling movie/video content.
Default: href (hyperlink to movie file).
Other options: media9, movie15, multimedia (Beamer's \movie command).
--latex_movie_controls=...
Specify control panel for movies. Default: on. Other options: off.
--latex_external_movie_viewer
Allow external movie viewer for movie15 package.
--latex_fancy_header
Typesetting of headers on each page:
If article: section name to the left and page number to the right
on even page numbers, the other way around on odd page numners.
If book: section name to the left and page numner to the right
on even page numbers, chapter name to the right and page number to
the left on odd page numbers.
--latex_section_headings=...
Typesetting of title/section/subsection headings: std (default): standard LaTeX,
blue: gray blue color,
strongblue: stronger blue color,
gray: white text on gray background, fit to heading width,
gray-wide: white text on gray background, wide as the textwidth.
--latex_colored_table_rows=...
Colors on every two line in tables: no (default), gray, blue.
--latex_line_numbers
Include line numbers for the running text (only active if there
are inline comments.
--latex_todonotes
Use the todonotes package to typeset inline comments. Gives colored bubbles in the margin for small inline comments and
in the text for larger comments.
--latex_double_spacing
```

```
inline comments).
--latex_labels_in_margin
Print equation, section and other LaTeX labels in the margin.
--latex_index_in_margin
Place entries in the index also in the margin.
--latex_preamble=...
User-provided LaTeX preamble file, either complete or additions
to the doconce-generated preamble.
--latex_no_program_footnotelink
If --device=paper, this option removes footnotes with links to
computer programs.
--latex_admon=...
Type of admonition in LaTeX:
colors1:
(inspired by the NumPy User Guide) applies different colors
for the different admons with an embedded icon,
colors2:
like 'colors1' but the text is wrapped around the icon,
rounded gray boxes with a optional title and no icon (default),
graybox2:
box with square corners, gray background, and narrower
than mdfbox, if code it reduces to something like mdfbox
(mdframed based); the summary admon is in case of A4 format
only half of the text width with text wrapped around
(effective for proposals and articles),
grayicon:
box with gray icons and a default light gray background,
yellowicon:
box yellow icons and a default light yellow background,
paragraph: plain paragraph with boldface heading.
--latex_admon_color=...
The color to be used as background in admonitions.
Either rgb tuple or saturated color a la yellow!5:
  --latex_admon_color=0.1,0.1,0.4
 '--latex_admon_color=yellow!5'
(note the quotes, needed for bash, in the latter example)
```

Sets the LaTeX linespacing to 1.5 (only active if there are

```
--latex_admon_title_no_period
Do not add a period at the end of admon titles in LaTeX if
it is missing.
--latex_admon_envir_map=...
Mapping of code envirs to new envir names inside admons
(e.g., to get a different code typesetting inside admons).
If a number, say 2, as in --latex_admon_envir_map=2,
an envir like pycod gets the number appended:

pycod2. Otherwise it must be a mapping for each envir:

--latex_admon_envir_map=pycod-pycod_yellow,fpro-fpro2
(from-to, from-to, ... syntax).
--latex_exercise_numbering=...
absolute: exercises numbered as 1, 2, ...
chapter: exercises numbered as 1.1, 1.2, ..., 3.1, 3.2, etc.
          with a chapter prefix.
--latex_subex_header_postfix=...
Default: ).
Gives headers a), b), etc. Can be set to period, colon, etc.
--xelatex
Use xelatex instead of latex/pdflatex.
--latex_double_hyphen
Replace single dash - by double dash -- in LaTeX output. Somewhat intelligent, but may give unwanted edits. Use with great care!
--verbose
Write out all OS commands run by doconce.
--examples_as_exercises
Treat examples of the form "==== Example: ..."
as in exercise environments.
--without_solutions
Leave out solution environments from exercises.
--without_answers
```

Leave out answer environments from exercises.

```
Leave out hints from exercises.
--wordpress
Make HTML output for wordpress.com pages.
--tables2csv
Write each table to a CSV file table_X.csv,
where X is the table number (autonumbered in according to
appearance in the DocOnce source file).
--sections_up
Upgrade all sections: sections to chapters, subsections
to sections, etc.
--sections_down
Downgrade all sections: chapters to sections, sections
to subsections, etc.
--os_prompt=...
Terminal prompt in output from running OS commands (the @@@OSCMD instruction). None or empty: no prompt, just the command; nocmd: no command, just the output. Default is "Terminal>".
--code_prefix=...
Prefix all @@@CODE imports with some path.
--figure_prefix=...
Prefix all figure filenames with, e.g., an URL.
--movie_prefix=...
Prefix all movie filenames with, e.g., an URL.
--no_mp4_webm_ogg_alternatives
Use just the specified (.mp4, .webm, .ogg) movie file; do not allow alternatives in HTML5 video tag.
Used if the just the specified movie format should be played.
--handout
Makes slides output suited for printing.
--urlcheck
```

--without\_hints

Check that all URLs referred to in the document are valid. --short\_title=... Short version of the document's title. Allow Markdown (and some Extended Markdown) syntax as input. --md2do\_output=... Dump to file the Doconce code arising from converting from Markdown. Default value is None (no dump). Any filename can be specified: --md2do\_output=myfile.do.txt --github\_md Turn on github-flavored-markdown dialect of the pandoc translator --strapdown Wrap Markdown output in HTML header/footer such that the output file (renamed as .html) can automatically be rendered as an HTML via strapdownjs.com technology. Combine with --github\_md for richer output. Styles are set with --bootwatch\_theme=cyborg (for instance). --strict\_markdown\_output Ensure strict/basic Markdown as output. --multimarkdown\_output Allow MultiMarkdown as output. --quiz\_question\_prefix=...  ${\tt Prefix/title\ before\ question\ in\ quizzes.\ Default:\ "Question:".}$ Can also be set in square brackets for each individual question. ("Q: [] What is 1+1?" results in no prefix/title before the "What is 1+1?". --quiz\_choice\_prefix=... Prefix/title before choices in quizzes.
Default for HTML: "Choice", resulting in numbered choices
"Choice 1:", "Choice 2:", etc. A value with colon, period, or question mark (e.g., "Answer:") leaves out the numbering. Default for latex/pdflatex: letter or letter+checkbox. Other values: number, number+checkbox, number+circle, letter+circle, letter. The checkbox or circle is always omitted if answers or solutions are

```
included (i.e., if none of the --without_answers and
--without_solutions is set).
The choice prefix can also be set in square brackets for each
individual choice.
("Cr: [] Two"
results in no prefix/title before the the answer "Two".

--quiz_horizontal_rule=...
on (default): <hr>> before and after quiz in HTML. off: no <hr>>.
--rst_uio
Univ. of Oslo version of rst files for their Vortex system.

--rst_mathjax
Use raw HTML with MathJax for LaTeX mathematics in rst files.
--oneline_paragraphs
Combine paragraphs to one line (does not work well).
```

# 7 Installation of DocOnce and its Dependencies

Below, we explain the manual installation of all software that may be needed when working with DocOnce documents. The impatient way to install what is needed on a Debian-based Linux computer (running, e.g., Ubuntu) is to run the install\_doconce.sh (or install\_doconce.py) script.

## 7.1 DocOnce

DocOnce itself is pure Python code hosted at https://github.com/hplgit/doconce. Installation can be done via

```
sudo pip install -e git+https://github.com/hplgit/doconce#egg=doconce
# or if doconce is already installed
sudo pip install -e git+https://github.com/hplgit/doconce#egg=doconce --upgrade
```

Alternatively, one can run the standard procedure

```
git clone git@github.com:hplgit/doconce.git
cd doconce
sudo python setup.py install
cd ..
```

Since DocOnce is frequently updated, it is recommended to use the above procedure and whenever a problem occurs, make sure to update to the most recent version:

```
cd doconce
git pull origin master
sudo python setup.py install
```

## 7.2 Dependencies

Producing HTML documents, plain text, pandoc-extended Markdown, and wikis can be done without installing any other software. However, if you want other formats as output (LATEX, Sphinx, reStructuredText) and assisting utilities such as preprocesors, spellcheck, file differences, bibliographies, and so on, the software below must be installed.

To install all needed packages on a GNU/Linux Debian system, such as Ubuntu, you can jump to the script in Section 7.3.

**Preprocessors.** If you make use of the Preprocess preprocessor, this program must be installed:

```
svn checkout http://preprocess.googlecode.com/svn/trunk/ preprocess
cd preprocess
cd doconce
sudo python setup.py install
cd ..
```

A much more advanced alternative to Preprocess is Mako. Its installation is most conveniently done by pip,

```
pip install Mako
```

This command requires pip to be installed. On Debian Linux systems, such as Ubuntu, the installation is simply done by

```
sudo apt-get install python-pip
```

Alternatively, one can install from the pip source code.

Make can also be installed directly from source: download the tarball, pack it out, go to the directory and run the usual sudo python setup.py install.

**Image file handling.** Different output formats require different formats of image files. For example, PostScript or Encapuslated PostScript is required for latex output, while HTML needs JPEG, GIF, or PNG formats. DocOnce calls up programs from the ImageMagick suite for converting image files to a proper format if needed. The ImageMagick suite can be installed on all major platforms. On Debian Linux (including Ubuntu) systems one can simply write

```
sudo apt-get install imagemagick
```

The convenience program doconce <code>combine\_images</code>, for combining several images into one, will use <code>montage</code> and <code>convert</code> from ImageMagick and the <code>pdftk</code>, <code>pdfnup</code>, and <code>pdfcrop</code> programs from the <code>texlive-extra-utils</code> Debian package. The latter gets installed by

```
sudo apt-get install texlive-extra-utils
```

Automatic image conversion from EPS to PDF calls up  ${\tt epstopdf},$  which can be installed by

```
sudo apt-get install texlive-font-utils
```

**Spellcheck.** The utility doconce spellcheck applies the ispell program for spellcheck. On Debian (including Ubuntu) it is installed by

```
sudo apt-get install ispell
```

**Bibliography.** The Python package Publish is needed if you use a bibliography in your document. On the website, click on *Clone*, copy the command and run it:

```
hg clone https://bitbucket.org/logg/publish
```

Thereafter go to the publish directory and run the setup.py script for installing Publish:

```
cd publish
sudo python setup.py
```

Ptex2tex for LATEX Output. To make LATEX documents with very flexible choice of typesetting of verbatim code blocks you need ptex2tex, which is installed by

```
svn checkout http://ptex2tex.googlecode.com/svn/trunk/ ptex2tex
cd ptex2tex
sudo python setup.py install
```

It may happen that you need additional style files, you can run a script, cp2texmf.sh:

```
cd latex
sh cp2texmf.sh # copy stylefiles to ~/texmf directory
cd /
```

This script copies some special stylefiles that that ptex2tex potentially makes use of. Some more standard stylefiles are also needed. These are installed by

```
sudo apt-get install texlive
```

on Debian Linux (including Ubuntu) systems. TeXShop on Mac comes with the necessary stylefiles (if not, they can be found by googling and installed manually in the ~/texmf/tex/latex/misc directory).

Note that the doconce ptex2tex command, which needs no installation beyond DocOnce itself, can be used as a simpler alternative to the ptex2tex program.

The *minted* LaTeX style is offered by ptex2tex and doconce ptext2tex and popular among many users. This style requires the package Pygments to be installed. On Debian Linux,

```
sudo apt-get install python-pygments
```

Alternatively, the package can be installed manually:

```
hg clone ssh://hg@bitbucket.org/birkenfeld/pygments-main pygments cd pygments sudo python setup.py install
```

One can also do the simple

```
pip install sphinx
```

which also installs pygments.

If you use the minted style together with ptex2tex, you have to enable it by the -DMINTED command-line argument to ptex2tex. This is not necessary if you run the alternative doconce ptex2tex program.

All use of the minted style requires the -shell-escape command-line argument when running LATEX, i.e., latex -shell-escape or pdflatex -shell-escape.

Inline comments apply the todonotes LATEX package if the option --latex\_todonotes is given. The todonotes package requires several other packages: xcolor, ifthen, xkeyval, tikz, calc, graphicx, and setspace. The relevant Debian packages for installing all this are listed below.

LATEX packages. Many LATEX packages are potentially needed (depending on various preprocessor variables given to ptex2tex or doconce ptex2tex. The standard packages always included are relsize, epsfig, makeidx, setspace, color, amsmath, amsfonts, xcolor, bm, microtype, titlesec, and hyperref. The ptex2tex package (from ptex2tex) is also included, but removed again if doconce ptex2tex is run instead of the ptex2tex program, meaning that if you do not use ptex2tex, you do not need ptex2tex.sty. Optional packages that might be included are minted, fontspec, xunicode, inputenc, helvet, mathpazo, wrapfig, calc, ifthen, xkeyval, tikz, graphicx, setspace, shadow, disable, todonotes, lineno, xr, framed, mdframe, movie15, a4paper, and a6paper.

Relevant Debian packages that gives you all of these LATEX packages are

```
texlive
texlive-extra-utils
texlive-latex-extra
texlive-font-utils
```

On old Ubuntu 12.04 one has to do sudo add-apt-repository ppa:texlive-backports/ppa and sudo apt-get update first, or alternatively install these as well:

texlive-math-extra texlive-bibtex-extra texlive-xetex texlive-humanities texlive-pictures Alternatively, one may pull in texlive-full to get all available style files.

If you want to use the *anslistings* code environment with ptex2tex(.ptex2tex.cfg styles Python\_ANS, Python\_ANSt, Cpp\_ANS, etc.) or doconce ptex2tex (envir=ans or envir=ans:nt), you need the anslistings.sty file. It can be obtained from the ptex2tex source. It should get installed by the cp2texmf.sh script executed above.

reStructuredText (reST) Output. The rst output from DocOnce allows further transformation to LaTeX, HTML, XML, OpenOffice, and so on, through the docutils package. The installation of the most recent version can be done by

```
svn checkout \
  http://docutils.svn.sourceforge.net/svnroot/docutils/trunk/docutils
cd docutils
sudo python setup.py install
cd ..
```

#### The command

```
pip install sphinx
```

installs Docutils along with Sphinx and Pygments.

To use the OpenOffice suite you will typically on Debian systems install sudo apt-get install unovonv libreoffice libreoffice-dmaths

There is a possibility to create PDF files from reST documents using ReportLab instead of LaTeX. The enabling software is rst2pdf. Either download the tarball or clone the svn repository, go to the rst2pdf directory and run the usual sudo python setup.py install.

**Sphinx Output.** Output to sphinx requires of course the Sphinx software, installed by

```
hg clone https://bitbucket.org/birkenfeld/sphinx cd sphinx sudo python setup.py install cd
```

#### An alternative is

```
pip install sphinx
```

DocOnce comes with many Sphinx themes that are not part of the standard Sphinx source distribution. Some of these themes require additional Python/Sphinx modules to be installed:

- cloud and redcloud: https://bitbucket.org/ecollins/cloud\_sptheme
- bootstrap: https://github.com/ryan-roemer/sphinx-bootstrap-theme
- solarized: https://bitbucket.org/miiton/sphinxjp.themes.solarized

- impressjs: https://github.com/shkumagai/sphinxjp.themes.impressjs
- sagecellserver: https://github.com/kriskda/sphinx-sagecell

These must be downloaded or cloned, and setup.py must be run as shown above.

**Markdown and Pandoc Output.** The DocOnce format pandoc outputs the document in the Pandoc extended Markdown format, which via the pandoc program can be translated to a range of other formats. Installation of Pandoc, written in Haskell, is most easily done by

sudo apt-get install pandoc
on Debian (Ubuntu) systems.

**Epydoc Output.** When the output format is epydoc one needs that program too, installed by

svn co https://epydoc.svn.sourceforge.net/svnroot/epydoc/trunk/epydoc epydoc
cd epydoc
sudo make install
cd ...

**Remark.** Several of the packages above installed from source code are also available in Debian-based system through the apt-get install command. However, we recommend installation directly from the version control system repository as there might be important updates and bug fixes. For svn directories, go to the directory, run svn update, and then sudo python setup.py install. For Mercurial (hg) directories, go to the directory, run hg pull; hg update, and then sudo python setup.py install.

Analyzing file differences. The doconce diff file1 file prog command for illustrating differences between two files file1 and file2 using the program prog requires prog to be installed. By default, prog is difflib which comes with Python and is always present if you have DocOnce installed. Another choice, diff, should be available on all Unix/Linux systems. Other choices, their URL, and their sudo apt-get install command on Debian (Ubuntu) systems appear in the table below.

Program	URL	Debian/Ubuntu install
pdiff	a2ps wdiff	sudo apt-get install a2ps wdiff texlive-latex-extra texlive-latex-re
latexdiff	latexdiff	<pre>sudo apt-get install latexdiff</pre>
kdiff3	kdiff3	sudo apt-get install kdiff3
diffuse	diffuse	sudo apt-get install diffuse
xxdiff	xxdiff	<pre>sudo apt-get install xxdiff</pre>
meld	meld	sudo apt-get install meld
tkdiff.tcl	tkdiff	not in Debian

## 7.3 Quick Debian/Ubuntu Install

On Debian (including Ubuntu) systems, it is straightforward to install the long series of DocOnce dependencies:

```
# Version control systems
sudo apt-get install -y mercurial git subversion
sudo apt-get install -y idle ipython python-pip python-pdftools texinfo
# LaTeX
sudo apt-get install -y texlive texlive-extra-utils texlive-latex-extra texlive-math-extra texli
# or sudo apt-get install -y texlive-full # get everything
sudo apt-get install -y latexdiff auctex
# Image and movie tools
sudo apt-get install -y imagemagick netpbm mjpegtools pdftk giftrans gv libav-tools libavcodec-e
# Misc
sudo apt-get install -y ispell pandoc libreoffice unoconv libreoffice-dmaths curl a2ps wdiff mel
# More Python software
sudo pip install sphinx # install pygments and docutils too
sudo pip install mako
sudo pip install -e svn+http://preprocess.googlecode.com/svn/trunk#egg=preprocess
sudo pip install -e hg+https://bitbucket.org/logg/publish#egg=publish
sudo pip install -e hg+https://bitbucket.org/ecollins/cloud_sptheme#egg=cloud_sptheme
sudo pip install -e git+https://github.com/ryan-roemer/sphinx-bootstrap-theme#egg=sphinx-bootstrsudo pip install -e hg+https://bitbucket.org/miiton/sphinxjp.themes.solarized#egg=sphinxjp.themes
sudo pip install -e git+https://github.com/shkumagai/sphinxjp.themes.impressjs#egg=sphinxjp.themsudo pip install -e git+https://github.com/kriskda/sphinx-sagecell#egg=sphinx-sagecell sudo pip install -e svn+https://epydoc.svn.sourceforge.net/svnroot/epydoc/trunk/epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=epydoc#egg=e
# DocOnce itself
rm -rf srclib
                                      # put downloaded software in srclib
mkdir srclib
cd srclib
git clone git@github.com:hplgit/doconce.git
cd doconce
sudo python setup.py install -y
cd ../..
# Ptex2tex
cd srclib
svn checkout http://ptex2tex.googlecode.com/svn/trunk/ ptex2tex
cd ptex2tex
sudo python setup.py install -y
cd latex
sh cp2texmf.sh # copy stylefiles to ~/texmf directory
cd ../../..
```

# 8 Basic Parsing Ideas

The (parts of) files with computer code to be directly included in the document are first copied into verbatim blocks.

All verbatim and TeX blocks are removed and stored elsewhere to ensure that no formatting rules are not applied to these blocks.

The text is examined line by line for typesetting of lists, as well as handling of blank lines and comment lines. List parsing needs some awareness of the context. Each line is interpreted by a regular expression

```
(?P<indent> *(?P<listtype>[*o-] )? *)(?P<keyword>[^:]+?:)?(?P<text>.*)\s?
```

That is, a possible indent (which we measure), an optional list item identifier, optional space, optional words ended by colon, and optional text. All lines are of this form. However, some ordinary (non-list) lines may contain a colon, and then the keyword and text group must be added to get the line contents. Otherwise, the text group will be the line.

When lists are typeset, the text is examined for sections, paragraphs, title, author, date, plus all the inline tags for emphasized, boldface, and verbatim text. Plain substitutions based on regular expressions are used for this purpose.

The final step is to insert the code and TeX blocks again (these should be untouched and are therefore left out of the previous parsing).

It is important to keep the DocOnce format and parsing simple. When a new format is needed and this format is not obtained by a simple edit of the definition of existing formats, it might be better to convert the document to reST and then to XML, parse the XML and write out in the new format. When the DocOnce format is not sufficient to getting the layout you want, it is suggested to filter the document to another, more complex format, say reST or LATEX, and work further on the document in this format.

# 8.1 Typesetting of Function Arguments, Return Values, and Variables

#### Note.

This text is somewhat outdated. Now DocOnce supports NumPy-style docstrings. DocOnce can also do automatic references to Python documentation, say the 'math' module, but the link only works in case of Sphinx output. Typical syntax is

```
With :mod:'numpy' and the :func:'scipy.io.loadmat' function in the :mod:'scipy.io' module, we can ...
```

In case of ambiguous names, say a local math module, one can use python, numpy, scipy, and mpl (matplotlib) as prefix (as defined in conf.py, generated by the doconce sphinx\_dir):

```
With :func:'python:math.sin' instead of :func:'math.sin' ...
```

(More info in the Sphinx API doc by the author - under development.)

A Mako function can provide greater flexibility such that other formats than Sphinx can take advantage of such references (full URL can be built into the function, depending on the format).

As part of comments (or doc strings) in computer code one often wishes to explain what a function takes of arguments and what the return values are. Similarly, it is desired to document class, instance, and module variables. Such arguments/variables can be typeset as description lists of the form listed below and *placed at the end of the doc string*. Note that argument, keyword argument, return, instance variable, class variable, and module variable are the only legal keywords (descriptions) for the description list in this context. If the output format is Epytext (Epydoc) or Sphinx, such lists of arguments and variables are nicely formatted.

- argument x: x value (float),
  which must be a positive number.
- keyword argument tolerance: tolerance (float) for stopping the iterations.
- return: the root of the equation (float), if found, otherwise None.
- instance variable eta: surface elevation (array).
- class variable items: the total number of MyClass objects (int).
- module variable debug: True: debug mode is on; False: no debugging (bool variable).

The result depends on the output format: all formats except Epytext and Sphinx just typeset the list as a list with keywords.

module variable x: x value (float), which must be a positive number.

module variable tolerance: tolerance (float) for stopping the iterations.

## 9 References

#### Note.

The references below are just for illustrating and testing citation syntax and not references for explaining parts of the text.

## References

- [1] D. E. Knuth. Theory and practice. EATCS Bulletin, 27:14-21, 1985.
- [2] H. P. Langtangen. A document for testing DocOnce. http://hplgit.github.io/doconce/test/demo\_testdoc.html.
- [3] H. P. Langtangen. The FEMDEQS program system. Research report in mechanics, Department of Mathematics, University of Oslo, 1989.

- [4] H. P. Langtangen. Numerical solution of first passage problems in random vibrations. *SIAM Journal of Scientific and Statistical Computing*, 15:997–996, 1994.
- [5] H. P. Langtangen, K.-A. Mardal, and R. Winther. Numerical methods for incompressible viscous flow. *Advances in Water Resources*, 25:1125–1146, 2002.
- [6] H. P. Langtangen and G. Pedersen. Propagation of large destructive waves. *International Journal of Applied Mechanics and Engineering*, 7(1):187–204, 2002.
- [7] K.-A. Mardal, G. W. Zumbusch, and H. P. Langtangen. Software tools for multigrid methods. In H. P. Langtangen and A. Tveito, editors, *Advanced Topics in Computational Partial Differential Equations – Numerical Methods and Diffpack Programming*, Lecture Notes in Computational Science and Engineering, pages 97–152. Springer, 2003.

## Index

```
AUTHOR keyword, 6
bibliography, 26
boldface words, 14
citations, 26
command-line options
    doconce format, 80
cross referencing, 22
DATE keyword, 6
demos, 4
doconce
    short explanation, 1
doconce format options, 80
emphasized words, 14
footnotes, 17
headlines, 7
index, 26
inline comments, 14
inline tagging, 14
labels, 22
nbsp, 14
non-breaking space, 14, 17
quotation syntax, 17
references, 22
section headings, 7
TITLE keyword, 6
TOC keyword, 7
verbatim text, 14
```