# hyperSpec Plotting functions

Claudia Beleites <chemometrie@beleites.de>
DIA Raman Spectroscopy Group, University of Trieste/Italy
Spectroscopy · Imaging, IPHT, Jena/Germany

November 18, 2016

# Reproducing the Examples in this Vignette

All spectra used in this manual are installed automatically with *hyperSpec*. Note that some definitions are executed in <code>vignette.defs</code>, and others invisibly at the beginning of the file in order to have the code as similar as possible to interactive sessions.

#### Contents

1	Predefined functions	2
2	Arguments for plot	4
3	Spectra 3.1 Stacked spectra	<b>7</b> 9
4	Calibration Plots, (Depth) Profiles, and Time Series Plots4.1Calibration plots4.2Time series and other Plots of the Type Intensity-over-Something	11 11 12
5	Levelplot	13
6	Spectra Matrix	13
7	False-Colour Maps: plotmap	15
8	3D plots (with rgl)	18
9	Using ggplot2 with hyperSpec objects	19
10	Troubleshooting 10.1 No output is produced	<b>20</b> 20
11	Interactive Graphics  11.1 spc.identify: finding out wavelength, intensity and spectrum	21

# Suggested Packages

latticeExtra: available deldir: available rgl: available ggplot2: available

In addition tripack, and latticist are mentioned, but not used in this vignette.

# **Preliminary Calculations**

For some plots of the chondro dataset, the pre-processed spectra and their cluster averages  $\pm$  one standard deviation are more suitable:

```
> chondro.preproc <- chondro - spc.fit.poly.below (chondro)
Fitting with npts.min = 15
> chondro.preproc <- chondro.preproc / rowMeans (chondro)
> chondro.preproc <- chondro.preproc - quantile (chondro, 0.05)
> cluster.cols <- c ("dark blue", "orange", "#C02020")
> cluster.meansd <- aggregate (chondro.preproc, chondro$clusters, mean_pm_sd)
> cluster.means <- aggregate (chondro.preproc, chondro$clusters, mean)</pre>
```

For details about the pre-processing, please refer to the example work flow in vignette ("chondro"), or the help? chondro.

#### 1 Predefined functions

hyperSpec comes with 6 major predefined plotting functions.

plot main switchyard for most plotting tasks

levelplot hyperSpec has a method for lattice[?] function levelplot

plotspc plots spectra

plotmat plots the spectra matrix

plotc calibration plot, time series, depth profile

plotc is a *lattice* function

plotmap more specialized version of levelplot for map or image plots.

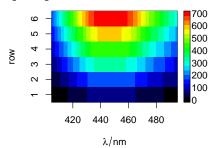
 ${\tt plotmap}$  is a lattice function

plotvoronoi more specialized version of plotmap that produces Voronoi tesselations.

plotvoronoi is a lattice function

plotmap, plotvoronoi, and levelplot are *lattice* functions. Therefore, in loops, functions, Sweave chunks, etc. the lattice object needs to be printed explicitly by e.g. print (plotmap (object)) (R FAQ: Why do lattice/trellis graphics not work?).

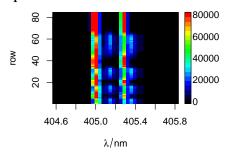
# plotspc



plots the spectra, i.e. the intensities \$spc over the wavelengths @wavelength.

> plotspc (flu)

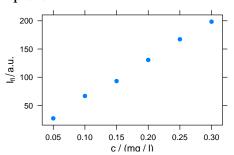
# plotmat



plots the spectra, i.e. the colour coded intensities  $\protect\$  over the wavelengths  $\protect\$  and the row number.

> plotmat (flu)

# plotc

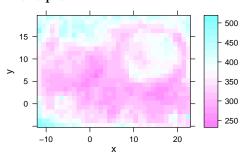


plots an intensity over a single other data column, e.g.

- $\bullet$  calibration
- time series
- depth profile

> plotc (flu)

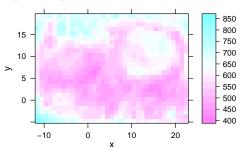
# levelplot



plots a false colour map, defined by a formula.

> levelplot (spc ~ x \* y, chondro, aspect = "iso")
Warning: Only first wavelength is used for plotting

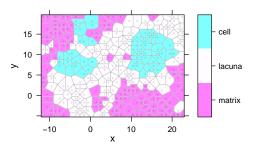
# plotmap



plotmap is a specialized version of level plot. It uses a single value (e. g. average intensity or cluster membership) over two data columns (default x and y

> plotmap (chondro)

# plotvoronoi

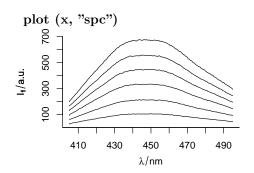


plotmap is a specialized version of level plot. It uses a single value (e.g. average intensity or cluster membership) over two data columns (default \$x and \$y)

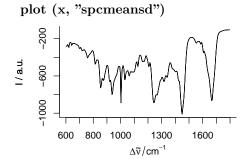
> plotvoronoi (sample (chondro, 300), clusters  $\tilde{x} * y$ )

# 2 Arguments for plot

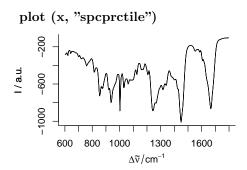
hyperSpec's plot method uses its second argument to determine which of the specialized plots to produce. This allows some handy abbreviations. All further arguments are handed over to the function actually producing the plot.



is equivalent to plotspc (flu)
> plot (flu, "spc")

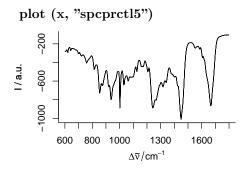


plots mean spectrum ± 1 standard deviation
> plot (chondro.preproc, "spcmeansd")

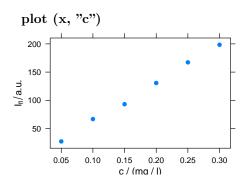


plots median,  $16^{\rm th}$  and  $84^{\rm th}$  percentile for each wavelength. For Gaussian distributed data,  $16^{\rm th}$ ,  $50^{\rm th}$  and  $84^{\rm th}$  percentile are equal to mean  $\pm$  standard deviation. Spectroscopic data frequently are not Gaussian distributed. The percentiles give a better idea of the true distribution. They are also less sensitive to outliers.

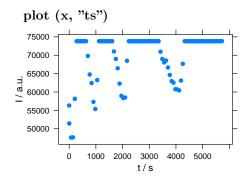
> plot (chondro.preproc, "spcprctile")



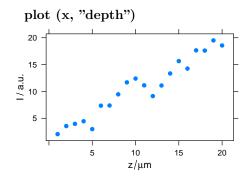
like "spcprct1" plus  $5^{\rm th}$  and  $95^{\rm th}$ percentile. > plot (chondro.preproc, "spcprct15")

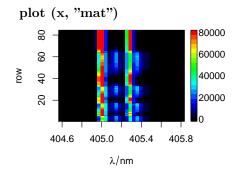


> plot (flu, "c")
is equivalent to plotc (flu)

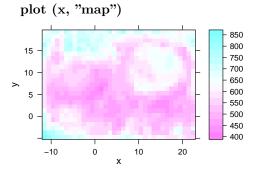


plots a time series plot
> plot (laser [,, 405], "ts")
equivalent to plotc (laser, spc ~ t)





plots the spectra matrix.
> plot (laser, "mat")
Equivalent to
> plotmat (laser)
A lattice alternative is:
> levelplot (spc ~ .wavelength \* .row, laser)



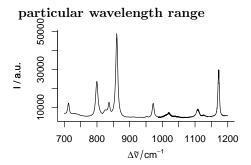
is equivalent to plotmap (chondro)
> plot (chondro, "map")

#### plot (x, "voronoi") > plot (sample (chondro, 300), "voronoi") 15 700 See plotvoronoi 650 10 600 5 550 500 0 450 -10 10 20

# 3 Spectra

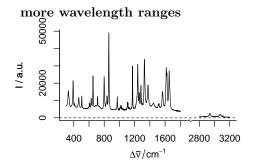
plotspc

 ${\tt plotspc}$  offers a variety of parameters for customized plots. To plot  $\dots$ 



if only one wavelength range is needed, the extract command
(see vignette ("introduction")) is handiest:
> plotspc (paracetamol [,, 700 ~ 1200])

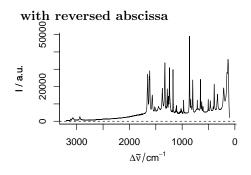
wavelengths. If wl.range already contains indices use
wl.index = TRUE.



use wl.range = list (600 ~ 1800, 2800 ~ 3100). Cut the wavelength axis appropriately with xoffset = 750 > plotspc (paracetamol,

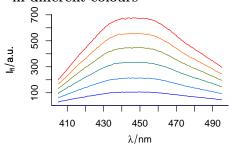
+ wl.range = c (300 ~ 1800, 2800 ~ max), + xoffset = 750)

If available, the package plotrix[1] is used to produce the cut mark.



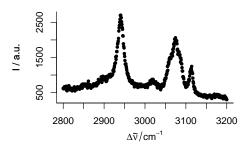
use wl.reverse = TRUE
> plotspc (paracetamol, wl.reverse = TRUE )

# in different colours

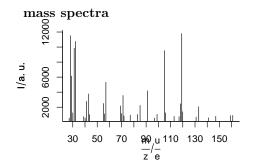


use col = vector.of.colours
> plotspc (flu, col = matlab.dark.palette (6))

# dots instead of lines

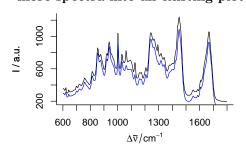


use lines.args = list (pch = 20, type = "p")
> plotspc (paracetamol [,, 2800 ~ 3200],
+ lines.args = list (pch = 20, type = "p"))



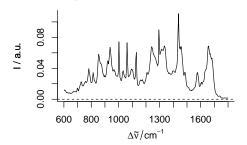
use lines.args = list (type = "h")
> plot (barbiturates [[1]], lines.args = list (type = "h"))

# more spectra into an existing plot



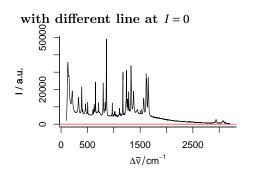
use add = TRUE
> plotspc (chondro [ 30,,])
> plotspc (chondro [300,,], add = TRUE, col = "blue")

# **Summary characteristics**



func may be used to calculate summary characteristics prior to plotting. To plot e.g. the standard deviation of the spectra, use:

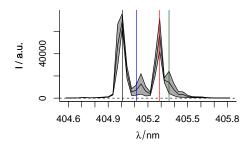
> plotspc (chondro.preproc, func = sd)



zero line takes a list with parameters to  ${\tt abline}, \, {\tt NA} \, \, {\tt suppresses}$  the line.

> plotspc (paracetamol,
+ zeroline = list (col = "red"))

```
adding to a spectra plot
```



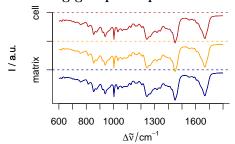
plotspc uses base graphics. After plotting the spectra, more content may be added to the graphic by abline, lines, points, etc.

> plot (laser, "spcmeansd")
> abline (v = c(405.0063, 405.1121, 405.2885, 405.3591),
+ col = c("black", "blue", "red", "darkgreen"))

# 3.1 Stacked spectra

# stacked "" 600 800 1000 1300 1600 Δῦ/cm<sup>-1</sup>

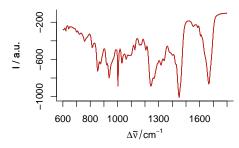
#### Stacking groups of spectra



The spectra to be stacked can be grouped: stacked = factor. Alternatively, the name of the grouping extra data column can be used:

```
> plot (cluster.meansd,
+ stacked = ".aggregate",
+ fill = ".aggregate",
+ col = cluster.cols)
```

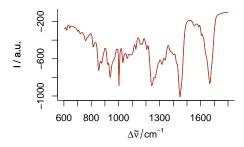
# Manually giving yoffset



Stacking values can also be given manually as numeric values in yoffset:

```
> plotspc (cluster.meansd,
+ yoffset = rep (0:2, each = 3),
+ col = rep (cluster.cols, each = 3))
```

#### Dense stacking



```
To obtain a denser stacking:
> yoffsets <- apply (cluster.means [[]], 2, diff)
> yoffsets <- apply (yoffsets, 1, min)
> plot (cluster.means, yoffset = c (0, cumsum (yoffsets)),
+ col = cluster.cols)
```

#### Elaborate example

plotspc allows fine grained customization of almost all aspects of the plot. This is possible by

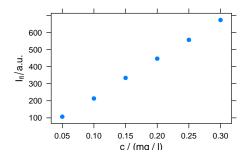
giving arguments to the functions that actually perform the plotting plot for setting up the plot area, lines for the plotting of the lines, axis for the axes, etc. The arguments for these functions should be given in lists as plot.args, lines.args, axis.args, etc.

# 4 Calibration Plots, (Depth) Profiles, and Time Series Plots

plotc

# 4.1 Calibration plots

# Intensities over concentration

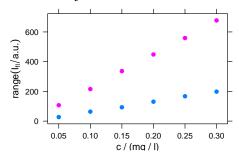


Plotting the Intensities of one wavelength over the concentration for univariate calibration:

> plotc (flu [,, 450])

The default is to use the first intensity only.

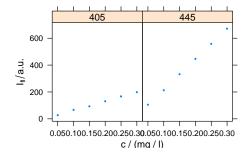
# Summary Intensities over concentration



A function to compute a summary of the intensities before drawing can be used:

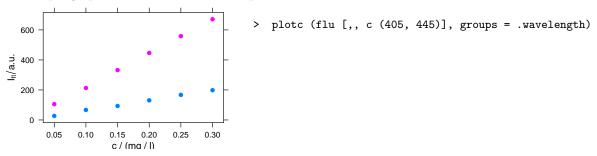
> plotc (flu, func = range, groups = .wavelength)
If func returns more than one value, the different results are accessible by .wavelength.

# Conditioning: plotting more traces separately

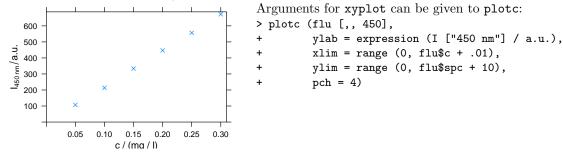


> plotc (flu [,, c (405, 445)], spc ~ c | .wavelength,
+ cex = .3, scales = list (alternating = c(1, 1)))

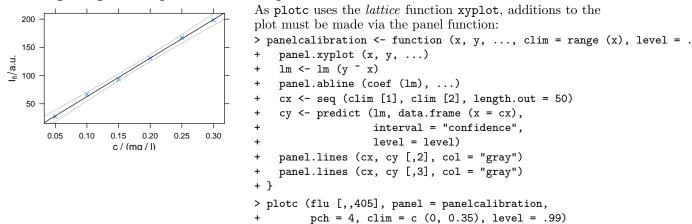
# Grouping: plot more traces in one panel



# Changing Axis Labels (and other parameters)

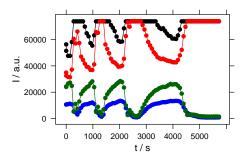


#### Adding things to the plot: customized panel function



# 4.2 Time series and other Plots of the Type Intensity-over-Something

#### Abscissae other than c



Other abscissae may be specified by explicitly giving the model formula:

# 5 Levelplot

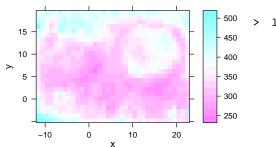
hyperSpec's levelplot can use two special column names:

.wavelength for the wavelengths

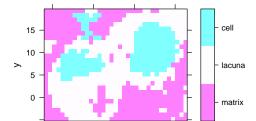
.row for the row index (i. e. spectrum number) in the data

Besides that, it behaves exactly like level plot. Particularly, the data is given as the second argument:

# levelplot



> levelplot (spc ~ x \* y, chondro)



10

20

If the colour-coded value is a factor, the display is adjusted to this fact:

> levelplot (clusters ~ x \* y, chondro)

# 6 Spectra Matrix

-10

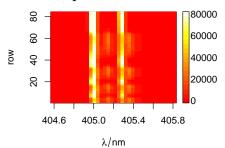
factors as z

It is often useful to plot the spectra against an additional coordinate, e.g. the time for time series, the depth for depth profiles, etc.

This can be done by plot (object, "mat"). The actual plotting is done by image, but levelplot

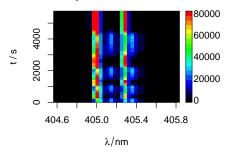
can produce spectra matrix plots as well and these plots can be grouped or conditioned.

# different palette



```
> plot (laser, "mat", col = heat.colors (20))
is the same as
> plotmat (laser, col = heat.colors (20))
```

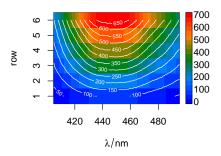
# different y axis



Using a different extra data column for the y axis: > plotmat (laser, y = "t") alternatively, y values and axis label can be given separately.

> plotmat (laser, y = laser\$t, ylab = labels (laser, "t"))

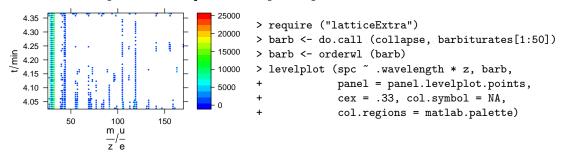
# contour lines



Contour lines may be added:

> plotmat (flu, col = matlab.dark.palette (20))
> plotmat (flu, col = "white",
+ contour = TRUE, add = TRUE)

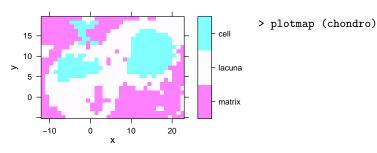
# colour-coded points: levelplot with special panel function



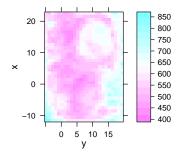
# 7 False-Colour Maps: plotmap

plotmap is a specialized version of levelplot. The spectral intensities may be summarized by a function before plotting (default: mean). The same scale is used for x and y axes (aspect = "iso").

# plotting map

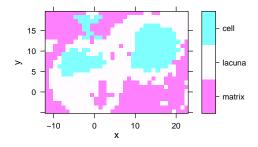


# plotting maps with other than x and y



specify the colour-coded variable, abscissa and ordinate as
formula: colour.coded ~ abscissa \* ordinate
> plotmap (chondro, spc ~ y \* x)

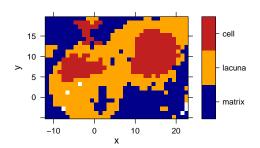
# colour-coded factors



> plotmap (chondro, clusters ~ x \* y)

If the colour-coded variable is a factor, each level gets its own colour, and the legend is labeled accordingly.

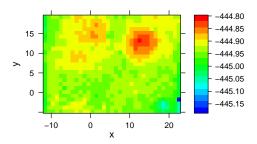
# different palette



To plot with a different palette, use trellis.args = list (col.regions = palette).

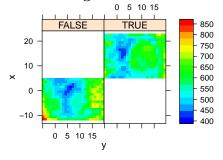
> print (plotmap (chondro, clusters ~ x \* y,
+ col.regions = cluster.cols))

# defined wavelengths



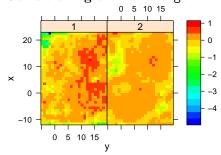
To plot a map of the average intensity at particular wavelengths use extraction:

# Conditioning



```
> plotmap (chondro,
+            spc ~ y * x | x > 5,
+            col.regions = matlab.palette(20))
```

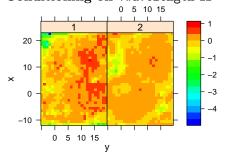
# Conditioning on .wavelength



plotmap automatically applies the function in *func* before plotting. This defaults to the mean. In order to suppress this, use func = NULL. This allows conditioning on the wavelengths.

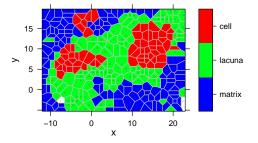
To plot e.g. the first two score maps of a principal component analysis:

# Conditioning on .wavelength ${\bf II}$



Alternatively, use levelplot directly:

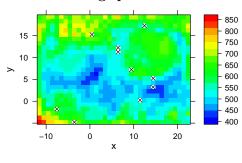
# Voronoi plot



```
> plotvoronoi (sample (chondro, 300), clusters ~ x * y,
+ col.regions = matlab.palette(20))
```

Voronoi uses panel.voronoi from latticeExtra[2]. The tesselation is calculated by default using deldir[3], but tripack[4] can also be used. tripack seems to faster in general, but may "hang" with certain data sets (particularly regular grids with missing spectra as in this example). Furthermore, it is not FOSS (free and open source software), so users are kindly asked to review tripack's license before using it.

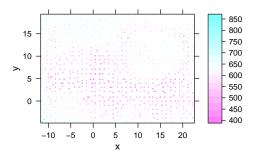
# Mark missing spectra



If the spectra come from a rectangular grid, missing positions can be marked with this panel function:

# Unevenly spaced measurement grid

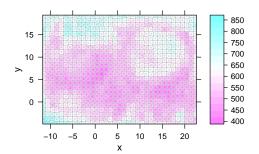
The panel function used by plotmap defaults to panel.levelplot.raster which assumes an evenly spaced measurement grid. Even if the spectra are measured on a nominally evenly spaced grid, the actual stage position may be slightly varying due to positioning inaccuracy and some manufacturers (e.g. Kaiser) record the position reported by the stage rather than the position requested by the stage control.



This leads to weird looking output with holes, and possibly wrong columns:

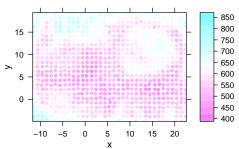
#### > plotmap (uneven)

The symptom of this situation are warnings about values in x and/or y not being equispaced; and that the output therefore may be wrong.



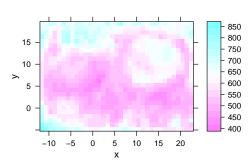
One possibility to obtain a correct map is using plotvoronoi instead which will construct a mosaic-like image with the respective "pixel" areas being centered around the actually recorded \$x\$ and \$y\$ position:

> plotvoronoi (uneven)



Another possibility that underlines a point shape of the measurements is switching to latticeExtra::panel.levelplot.points:

> plotmap (uneven, panel = panel.levelplot.points,
+ cex = 0.75, col.symbol=NA)



Alternatively, the measurement raster positions can be rounded to their nominal raster, e.g.:

> rx <- makeraster (uneven\$x, start = -11.55, d = 1, tol = 0.3)

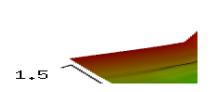
> uneven\$x <- rx\$x</pre>

> ry <- makeraster (uneven\$y, start = -4.77, d = 1, tol = 0.3)

> uneven\$y <- ry\$x

> plotmap (uneven)

8 3D plots (with rgl)



rgl[5] offers fast 3d plotting in R. As rgl's axis annotations are sometimes awkward, they may better be set manually:

> laser <- laser [,,404.8 ~ 405.6] / 10000

> laser\$t <- laser\$t / 3600

> cols <- rep (matlab.palette (nrow (laser)), nwl (laser))

surface3d (y = wl (laser), x = laser\$t,

z = laser\$spc, col = cols)

> aspect3d (c (1, 1, 0.25))

> axes3d (c ('x+-', 'y--', 'z--'))

> axes3d ('y--', nticks = 25, labels= FALSE)

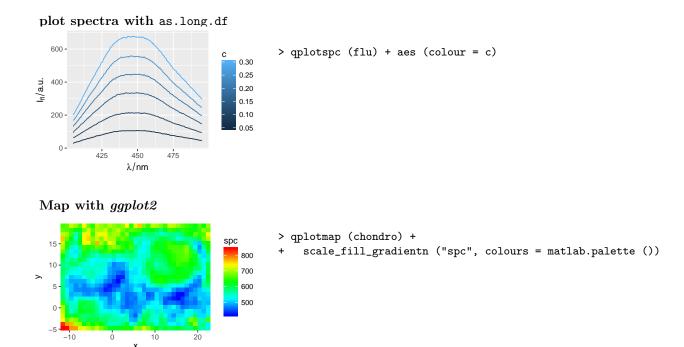
> mtext3d ("t / h", 'x+-', line = 2.5)

> mtext3d ("lambda / nm", 'y--', line = 2.5)

> mtext3d ("I / a.u.", edge = 'z--', line = 2.5)

# 9 Using ggplot2 with hyperSpec objects

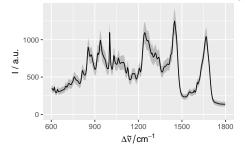
hyperSpec objects do not yet directly support plotting with ggplot2 [6]. Nevertheless, ggplot2 graphics can easily be obtained, and qplot\* equivalents to plotspc and plotmap are defined:



The two special columns .wavelength and .rownames contain the wavelength axis and allow to distinguish the spectra.

For more general plotting, as.long.df transforms a hyperSpec object into a long-form data.frame that is suitable for qplot, while as.t.df produces a data.frame where each spectrum is one column, and an additional first column gives the wavelength (see "plotting mean  $\pm$  sd" below for an example). Long data.frames can be very memory consuming as they are of size  $nrow \cdot nwl \times (ncol + 2)$  with respect to the dimensions of the hyperSpec object. Thus, e.g. the chondro data set (2 MB) as hyperSpec object) needs 24 MB as long-format data.frame. It is therefore highly recommended to calculate the particular data to be plotted beforehand.

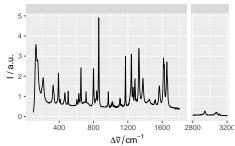
# Mean $\pm$ standard deviation with ggplot2



```
> qplotspc (mean (chondro)) +
+ geom_ribbon (aes (ymin = mean + sd,
+ ymax = mean - sd,
+ y = 0, group = NA),
+ alpha = 0.25,
+ data = as.t.df (mean_sd (chondro)))
```

Note that qpotspc specifies aesthetics y = spc and groups = .rownames, which do not have corresponding columns in the data.frame returned by as.t.df. These aesthetics must therefore be set manually in the aesthetics definition in geom\_ribbon (or any other geom\_ that uses as.t.df). Otherwise, errors occur that object spc (and/or .rownames) cannot be found.

# Cut spectra with gaplot2



# 10 Troubleshooting

#### 10.1 No output is produced

plotmap, plotvoronoi, levelplot, and plotc use *lattice* functions. Therefore, in loops, functions, Sweave chunks, etc. the lattice object needs to be printed explicitly by print (plotmap (object)) (R FAQ: Why do lattice/trellis graphics not work?). The same holds for *qqplot2* graphics.

For suggestions how the lattice functions can be redefined so that the result is printed without external print command, see the file vignettes.defs.

#### 11 Interactive Graphics

hyperSpec offers basic interaction, spc.identify for spectra plots, and map.identify and map.sel.poly for maps. The first two identify points in spectra plots and map plots, respectively. map.sel.poly selects the part of a hyperSpec object that lies inside the user defined polygon.

#### 11.1 spc.identify: finding out wavelength, intensity and spectrum

spc.identify allows to measure points in graphics produced by plotspc. It works correctly with reversed and cut wavelength axes.

```
> spc.identify (plotspc (paracetamol, wl.range = c (600 ~ 1800, 2800 ~ 3200), xoffset = 800))
```

The result is a data frame with the indices of the spectra, the wavelength, and its intensity.

# 11.2 map.identify: finding a spectrum in a map plot

```
map.identify returns the spectra indices of the clicked points.
> map.identify (chondro)
```

#### 11.3 map.sel.poly: selecting spectra inside a polygon in a map plot

map.sel.poly returns a logical indicating which spectra are inside the polygon drawn by the user:
> map.sel.poly (chondro)

#### 11.4 Related functions provided by base graphics and lattice

For base graphics (as produced by plotspc), locator may be useful as well. It returns the clicked coordinates. Note that these are *not* transformed according to xoffset & Co.

For lattice graphics, grid.locator may be used instead. If it is not called in the panel function, a preceding call to trellis.focus is needed:

```
> plot (laser, "mat")
> trellis.focus ()
> grid.locator ()
```

identify (or panel.identify for lattice graphics) allows to identify points of the plot directly. Note that the returned indices correspond to the plotted object.

#### References

- [1] Lemon J. Plotrix: a package in the red light district of r. R-News, 6(4):8-12, 2006.
- [2] Deepayan Sarkar and Felix Andrews. latticeExtra: Extra Graphical Utilities Based on Lattice, 2016. URL https://CRAN.R-project.org/package=latticeExtra. R package version 0.6-28.
- [3] Rolf Turner. deldir: Delaunay Triangulation and Dirichlet (Voronoi) Tessellation, 2016. URL https://CRAN.R-project.org/package=deldir. R package version 0.1-12.
- [4] Fortran code by R. J. Renka. R functions by Albrecht Gebhardt. With contributions from Stephen Eglen <stephen@anc.ed.ac.uk>, Sergei Zuyev, and Denis White. tripack: Triangulation of Irregularly Spaced Data, 2015. URL https://CRAN.R-project.org/package=tripack. R package version 1.3-7.
- [5] Daniel Adler, Duncan Murdoch, and others. rgl: 3D Visualization Using OpenGL, 2016. URL https://CRAN.R-project.org/package=rgl. R package version 0.96.0.
- [6] Hadley Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL http://ggplot2.org.

#### **Session Info**

```
[,1]
sysname "Linux"
release "4.4.0-21-generic"
version "#37-Ubuntu SMP Mon Apr 18 18:33:37 UTC 2016"
nodename "cb-T61p"
machine "x86_64"
login "unknown"
user "cb"
effective_user "cb"
```

R version 3.3.2 (2016-10-31)

Platform: x86\_64-pc-linux-gnu (64-bit) Running under: Ubuntu 16.04.1 LTS

locale:

[7] LC\_PAPER=de\_DE.UTF-8 LC\_NAME=C LC\_ADDRESS=C

[10] LC\_TELEPHONE=C LC\_MEASUREMENT=de\_DE.UTF-8 LC\_IDENTIFICATION=C

attached base packages:

[1] tools grid stats graphics grDevices utils datasets methods base

other attached packages:

[1] latticeExtra\_0.6-28 RColorBrewer\_1.1-2 baseline\_1.2-1 MASS\_7.3-45

[5] hyperSpec\_0.98-20161118 ggplot2\_2.2.0 lattice\_0.20-34

loaded via a namespace (and not attached):

[1] Rcpp\_0.12.7 magrittr\_1.5 knitr\_1.13 munsell\_0.4.3 svUnit\_0.7-12 [6] xtable\_1.8-2 colorspace\_1.3-0 R6\_2.1.2 stringr\_1.1.0 plyr\_1.8.4 [11] gtable\_0.2.0 plotrix\_3.6-2 deldir\_0.1-12 htmltools\_0.3.5 lazyeval\_0.2.0 [16] assertthat\_0.1 digest\_0.6.10 tibble\_1.2 shiny\_0.13.2 reshape2\_1.4.2 [21] htmlwidgets\_0.6 mime\_0.4 labeling\_0.3 stringi\_1.1.2 scales\_0.4.1