NEWS for R version 4.3.0 (2023-04-21)



CHANGES IN R 4.3.0

SIGNIFICANT USER-VISIBLE CHANGES:

- Calling `&&` or `||` with LHS or (if evaluated) RHS of length greater than one is now always an error, with a report of the form

  ```
  'length = 4' in coercion to 'logical(1)'
  ```

  Environment variable `_R_CHECK_LENGTH_1_LOGIC2_` no longer has any effect.

NEW FEATURES:

- The included BLAS sources have been updated to those shipped with LAPACK version 3.10.1. (This caused some platform-dependent changes to package check output.) And then to the sources from LAPACK version 3.11.0 (with changes only to double complex subroutines).
- The included LAPACK sources have been updated to include the four Fortran 90 routines rather than their Fortran 77 predecessors. This may give some different signs in SVDs or eigendecompositions.. (This completes the transition to LAPACK 3.10.x begun in R 4.2.0.)
- The LAPACK sources have been updated to version 3.11.0. (No new subroutines have been added, so this almost entirely bug fixes: Those fixes do affect some computations with `NaN`s, including R's `NA`.)
- The parser now signals *classed* errors, notably in case of the pipe operator `|>`. The error object and message now give line and column numbers, mostly as proposed and provided by Duncan Murdoch in PR#18328.
- `toeplitz()` is now generalized for asymmetric cases, with a `toeplitz2()` variant.
- `xy.coords()` and `xyz.coords()` and consequently, e.g., `plot(x,y, log = "y")` now signal a *classed* warning about negative values of y (where `log(.)` is `NA`). Such a warning can be specifically suppressed or caught otherwise.
- Regular expression functions now check more thoroughly whether their inputs are valid strings (in their encoding, e.g. in UTF-8).
- The performance of `grep()`, `sub()`, `gsub()` and `strsplit()` has been improved, particularly with `perl = TRUE` and `fixed = TRUE`. Use of `useBytes = TRUE` for performance reasons should no longer be needed and is discouraged: it may lead to incorrect results.

- `apropos()` gains an argument `dot_internals` which is used by the completion (`help(rcompgen)`) engine to also see **base** internals such as `.POSIXct()`.
- Support in `tools::Rdiff()` for comparing uncompressed PDF files is further reduced – see its help page.
- `qqplot(x, y, ...)` gains `conf.level` and `conf.args` arguments for computing and plotting a confidence band for the treatment function transforming the distribution of `x` into the distribution of `y` (Switzer, 1976, *Biometrika*). Contributed by Torsten Hothorn.
- Performance of `package_dependencies()` has been improved for cases when the number of dependencies is large.
- Strings newly created by `gsub()`, `sub()` and `strsplit()`, when any of the inputs is marked as `"bytes"`, are also marked as `"bytes"`. This reduces the risk of creating invalid strings and accidental substitution of bytes deemed invalid.
- Support for `readLines(encoding = "bytes")` has been added to allow processing special text files byte-by-byte, without creating invalid strings.
- `iconv(from = "")` now takes into account any declared encoding of the input elements and uses it in preference to the native encoding. This reduces the risk of accidental creation of invalid strings, particularly when different elements of the input have different encoding (including `"bytes"`).
- Package repositories in `getOption("repos")` are now initialized from the '`repositories`' file when **utils** is loaded (if not already set, e.g., in '`.Rprofile`'). (From a report and patch proposal by Gabriel Becker in PR#18405.)
- `compactPDF()` gets a `verbose` option.
- `type.convert()` and hence `read.table()` get new option `tryLogical = TRUE` with back compatible default. When set to false, converts `"F"` or `"T"` columns to character.
- Added new unit prefixes `"R"` and `"Q"` for abbreviating (unrealistically large) sizes beyond 10^{27} in `standard = "SI"`, thanks to Henrik Bengtsson's PR#18435.
- `as.data.frame()`'s default method now also works fine with atomic objects inheriting from classes such as `"roman"`, `"octmode"` and `"hexmode"`, such fulfilling the wish of PR#18421, by Benjamin Feakins.
- The `as.data.frame.vector()` utility now errors for wrong-length `row.names`. It warned for almost six years, with "Will be an error!".
- `sessionInfo()` now also contains `La_version()` and reports codepage and timezone when relevant, in both `print()` and `toLatex()` methods which also get new option `tzone` for displaying timezone information when `locale = FALSE`.
- New function `R_compiled_by()` reports the C and Fortran compilers used to build R, if known.
- `predict(<lm>, newdata = *)` no longer unnecessarily creates an `offset` of all `0`s.
- `solve()` for complex inputs now uses argument `tol` and by default checks for 'computational singularity' (as it long has done for numeric inputs).
- `predict(<rank-deficient lm>, newdata=*)` now obeys a new argument `rankdeficient`, with new default `"warnif"`, warning only if there are non-estimable cases in `newdata`. Other options include `rankdeficient = "NA"`, predicting `NA` for non-estimable `newdata` cases. This addresses PR#15072 by Russ Lenth and is based on his original proposal and discussions in PR#16158 also by David Firth and Elin Waring. Still somewhat experimental.
- `Rgui` console implementation now works better with the `NVDA` screen reader when the full blinking cursor is selected. The underlying improvements in cursor handling may help also other screen readers on Windows.

- The drop-field control in GraphApp can now be left with the TAB key and all controls can be navigated in the reverse order using the Shift+TAB key, improving accessibility of the `Rgui` configuration editor.
- `qnorm(<very large negative>, log.p=TRUE)` is now fully accurate (instead of to "only" minimally five digits).
- `demo(error.catching)` now also shows off `withWarnings()` and `tryCatchWEMs()`.
- As an experimental feature the placeholder `_` can now also be used in the `rhs` of a forward pipe `|>` expression as the first argument in an extraction call, such as `_$coef`. More generally, it can be used as the head of a chain of extractions, such as `_$coef[[2]]`.
- Spaces in the environment variable used to choose the R session's temporary directory (`TMPDIR`, `TMP` and `TEMP` are tried in turn) are now fatal. (On Windows the 'short path' version of the path is tried and used if that does not contain a space.)
- `all.equal.numeric()` gets a new optional switch `giveErr` to return the numeric error as attribute. Relatedly, `stopifnot(all.equal<some>(a, b, ..))` is as "smart" now, as `stopifnot(all.equal(....))` has been already, thus allowing customized `all.equal<Some>()` wrappers.
- R on Windows is now able to work with path names longer than 260 characters when these are enabled in the system (requires at least Windows 10 version 1607). Packages should be updated to work with long paths as well, instead of assuming `PATH_MAX` to be the maximum length. Custom front-ends and applications embedding R need to update their manifests if they wish to allow this feature. See https://blog.r-project.org/2023/03/07/path-length-limit-on-windows for more information.
- 'Object not found' and 'Missing argument' errors now give a more accurate error context. Patch provided by Lionel Henry in PR#18241.
- The `@` operator is now an S3 generic. Based on contributions by Tomasz Kalinowski in PR#18482.
- New generic `chooseOpsMethod()` provides a mechanism for objects to resolve cases where two suitable methods are found for an Ops Group Generic. This supports experimenting with alternative object systems. Based on contributions by Tomasz Kalinowski in PR#18484.
- `inherits(x, what)` now accepts values other than a simple character vector for argument `what`. A new generic, `nameOfClass()`, is called to resolve the class name from `what`. This supports experimenting with alternative object systems. Based on contributions by Tomasz Kalinowski in PR#18485.
- Detection of BLAS/LAPACK in use (`sessionInfo()`) with FlexiBLAS now reports the current backend.
- The `"data.frame"` method for `subset()` now warns about extraneous arguments, typically catching the use of '`=`' instead of '`==`' in the `subset` expression.
- Calling `a:b` when numeric `a` or `b` is longer than one may now be made into an error by setting environment variable `_R_CHECK_LENGTH_COLON_` to a true value, along the proposal in PR#18419 by Henrik Bengtsson.
- `density(x, weights = *)` now warns if automatic bandwidth selection happens without using `weights`; new optional `warnWbw` may suppress the warning. Prompted by Christoph Dalitz' PR#18490 and its discussants.
- `rm(list = *)` is faster and more readable thanks to Kevin Ushey's PR#18492.
- The `plot.lm()` function no longer produces a normal Q-Q plot for GLMs. Instead it plots a half-normal Q-Q plot of the absolute value of the standardized deviance residuals.

- The `print()` method for class `"summary.glm"` no longer shows summary statistics for the deviance residuals by default. Its optional argument `show.residuals` can be used to show them if required.
- The `tapply()` function now accepts a data frame as its `X` argument, and allows `INDEX` to be a formula in that case. `by.data.frame()` similarly allows `INDICES` to be a formula.
- The performance of `df[j] <- value` (including for missing `j`) and `write.table(df)` has been improved for data frames `df` with a large number of columns. (Thanks to Gabriel Becker's PR#18500, PR#18503 and discussants, prompted by a report from Toby Dylan Hocking on the R-devel mailing list.)
- The matrix multiply operator `%*%` is now an S3 generic, belonging to new group generic `matrixOps`. From Tomasz Kalinowski's contribution in PR#18483.
- New function `array2DF()` to convert arrays to data frames, particularly useful for the list arrays created by `tapply()`.

DATES and TIMES:

- On platforms where (non-UTC) datetimes before 1902 (or before 1900 as with system functions on recent macOS) are guessed by extrapolating time zones from 1902-2037, there is a warning at the first use of extrapolation in a session. (As all time zones post 2037 are extrapolation, we do not warn on those.)
- (Platforms using '`--with-internal-tzone`', including Windows and by default macOS). How years are printed in dates or date-times can be controlled by environment variable `R_PAD_YEARS_BY_ZERO`. The default remains to pad to 4 digits by zeroes, but setting value '`no`' gives no padding (as used by default by `glibc`).
- `strftime()` tries harder to determine the offset for the `"%z"` format, and succeeds on the mainstream R platforms.
- `strftime()` has a limit of 2048 bytes on the string produced – attempting to exceed this is an error. (Previously it silently truncated at 255 bytes.)
- `sessionInfo()` records (and by default prints) the system time zone as part of the locale information. Also, the source (system/internal) of the date-time conversion and printing functions.
- Objects of class `"POSIXlt"` created in this version of R always have 11 components: component `zone` is always set, and component `gmtoff` is set for times in UTC and usually set on the (almost all) platforms which have C-level support, otherwise is `NA`.
- There are comprehensive validity checks on the structure of objects of class `"POSIXlt"` when converting (including formatting and printing). (This avoids mis-conversions of hand-crafted objects.)
- There is some support for using the native date-time routines on macOS: this is only viable on recent versions (e.g. 12.6 and 13) and does get wrong some historical changes (before 1900, during WWII). Use of '`--with-internal-tzone`' remains the default.
- `as.POSIXct(<numeric>)` and `as.POSIXlt(.)` (without specifying `origin`) now work. So does `as.Date(<numeric>)`.
- `as.Date.POSIXct(., tz)` now treats several `tz` values, notably `"GMT"` as equivalent to `"UTC"`, proposed and improved by Michael Chirico and Joshua Ulrich in PR#17674.
- Experimental `balancePOSIXlt()` utility allows using "ragged" and or out-of-range `"POSIXlt"` objects more correctly, e.g., in subsetting and subassignments. Such objects are now documented.
  More experimentally, a `"POSIXlt"` object may have an attribute `"balanced"` indicating if it is known to be filled or fully balanced.

- Functions `axis.Date()` and `axis.POSIXct()` are rewritten to gain better default tick locations and better default formats by using `prettyDate()`. Thanks to Swetlana Herbrandt.
- The mapping of Windows' names for time zones to IANA's 'Olson' names has been updated. When ICU is available (it is by default), it is used to get a mapping for the current region set in Windows. This can be overridden by setting environment variable `TZ` to the desired Olson name — see `OlsonNames()` for those currently available.

GRAPHICS:

- The graphics engine version, `R_GE_version`, has been bumped to `16` and so packages that provide graphics devices should be reinstalled.
- The **grDevices** and **grid** packages have new functions for rendering typeset glyphs, primarily: `grDevices::glyphInfo()` and `grid::grid.glyph()`.
  Rendering of typeset glyphs is only supported so far on the Cairo-based graphics devices and on the `pdf()` and `quartz()` devices.
- The defined behaviour for `"clear"` and `"source"` compositing operators (via `grid::grid.group()`) has been changed (to align better with simple interpretation of original Porter-Duff definitions).
- Support for gradients, patterns, clipping paths, masks, groups, compositing operators, and affine transformations has been added to the `quartz()` device.

INSTALLATION on a UNIX-ALIKE:

- A system installation of generic LAPACK 3.10.0 or later will be preferred to the version in the R sources.
  `configure` option '`--with-lapack=no`' (equivalently '`--without-lapack`') forces compilation of the internal LAPACK sources.
  If '`--with-lapack`' is not specified, a system `liblapack` is looked for and used if it reports version 3.10.0 or later and does not contain BLAS routines.
  Packages using LAPACK will need to be reinstalled if this changes to using an external library.
- On '`aarch64`' Linux platforms using GCC, `configure` now defaults to '`-fPIC`' (instead of '`-fpic`'), as desired in PR#18326.
- `configure` now checks conversion of datetimes between `POSIXlt` and `POSIXct` around year 2020. Failure (which has been seen on platforms missing `tzdata`) is fatal.
- If `configure` option '`--with-valgrind-instrumentation`' is given value `1` or `2`, option '`--with-system-valgrind-headers`' is now the default and ignored (with a warning). It is highly recommended that the system headers are installed alongside `valgrind`: they are part of its packaging on some Linux distributions and packaged separately (e.g. in the '`valgrind-devel`' RPM) on others. `configure` will give a warning if they are not found.
  The system headers will be required in a future release of R to build with `valgrind` instrumentation.
- `libcurl` 8.x is now accepted by `configure`: despite a change in major version number it changes neither API nor ABI.

INSTALLATION on WINDOWS:

- The makefiles and installer scripts for Windows have been tailored to '`Rtools43`', an update of the '`Rtools42`' toolchain. It is based on `gcc` 12 and newer versions of MinGW-W64, binutils and libraries. At this time R-devel can still be built using Rtools42 without changes, but when R-devel is installed via the installer, it will by default look for Rtools43.

- Old make targets `rsync-extsoft` and 32-bit ones that are no longer needed have been removed.
- Default builds (including for packages) no longer select C99. Thus the C standard used is the default for the compiler, which for the toolchain in '`Rtools43`' is C17. (This is consistent with Unix builds.)

PACKAGE INSTALLATION:

- The default C++ standard has been changed to C++17 where available (which it is on all currently checked platforms): if not C++14 or C++11 is used if available otherwise C++ is not supported.
- `USE_FC_LEN_T` is the default: this uses the correct (compiler-dependent) prototypes for Fortran BLAS/LAPACK routines called from C/C++, and requires adjustment of many such calls – see 'Writing R Extensions' §6.6.1.
- There is initial support for C++23 as several compilers are now supporting `-std=c++23` or `-std=c++2b` or similar. As for C++20, there no additional `configure` checks for C++23 features beyond a check that the compiler reports a `__cplusplus` value greater than that in the C++20 standard. C++ feature tests should be used.
- There is support for a package to indicate the version of the C standard which should be used to compile it, and for the installing user to specify this. In most cases R defaults to the C compiler's default standard which is C17 (a 'bug-fix' of C11) – earlier versions of R or compilers may have defaulted to C99.
  Current options are:

  USE_C17 Use a standard that is at most C17. The intention is to allow legacy packages to still be installed when later C standards become the default, including packages using new keywords as identifiers or with K&R-style function declarations. This will use C17 if available, falling back to C11.

  USE_C90 Use the C90 (aka C89) standard. (As that standard did not require compilers to identify that version, all we can verify is that the compiler does not claim to be using a later standard. It may accept C99 features – for example `clang` accepts // to make comments.)

  USE_C99 Use the C99 standard. This should be rarely needed – it avoids the few new features of C11/C17 which can be useful if a package assumes them if C17 is specified and they are not implemented.

  USE_C23 Use C23 (or in future, later). Compiler/library support for C23 is still being implemented, but LLVM clang from 15.0.0 and the upcoming GCC 13 have quite extensive support.

  These can be specified as part of the '`SystemRequirements`' field in the package's '`DESCRIPTION`' file or *via* options '`--use-C17`' and so on of `R CMD INSTALL` and `R CMD SHLIB`.
  For further details see "Writing R Extensions" §1.2.5.
- (Windows) A '`src/Makefile.ucrt`' or '`src/Makefile.win`' file is now included after '*R_HOME*`/etc`*R_ARCH*`/Makeconf`' and so no longer needs to include that file itself. Installation of a package with such a file now uses a site '`Makevars`' file in the same way as a package with a '`src/Makevars.win`' file would.
- `configure` is now passed crucial variables such as `CC` and `CFLAGS` in its environment, as many packages were not setting them (as documented in 'Writing R Extensions' §1.2).
  This has most effect where `configure` is used to compile parts of the package – most often by `cmake` or `libtool` which obfuscate the actual compile commands used.
  Also used for `configure.win` and `configure.ucrt` on Windows.

FORTRAN FLAGS:

- The flag '`-fno-optimize-sibling-calls`' is no longer forced for `gfortran` 7 and later. It should no longer be needed now using 'hidden' character-length arguments when calling BLAS/LAPACK routines from C/C++ is the default even for packages. (Unless perhaps packages call Fortran code from C/C++ without using R's headers and without allowing for these arguments.)

C-LEVEL FACILITIES:

- The deprecated S-compatibility macros `DOUBLE_*` in '`R_ext/Constants.h`' (included by '`R.h`') have been removed.
- The deprecated legacy typedefs of `Sint` and `Sfloat` in header '`R.h`' are no longer defined, and that header no longer includes header '`limits.h`' from C nor '`climits`' from C++.
- New macro `CAD5R()` is provided in '`Rinternals.h`' and used in a few places in the R sources.
- ALTREP now supports `VECSXP` vectors. Contributed by Gabor Csardi in PR#17620.
- The `Rcomplex` definition (in header '`R_ext/Complex.h`') has been extended to prevent possible mis-compilation when interfacing with Fortran (PR#18430). The new definition causes compiler warnings with static initializers such as `{1, 2}`, which can be changed to `{.r=1,.i=2}`.
  Using the new definition from C++ depends on compiler extensions supporting C features that have not been incorporated into the C++ standards but are available in `g++` and `clang++`: this may result in C++ compiler warnings but these have been worked around for recent versions of common compilers (GCC, Apple/LLVM clang, Intel).
  It is intended to change the inclusion of header '`R_ext/Complex.h`' by other R headers, so C/C++ code files which make use of `Rcomplex` should include that header explicitly.

UTILITIES:

- `R CMD check` does more checking of package '`.Rd`' files, warning about invalid email addresses and (some) invalid URIs and noting empty '`\item`' labels in description lists.
- `R CMD check` now also reports problems when reading package news in md (file '`NEWS.md`') and (optionally) plain text (file '`NEWS`') formats.
- `_R_CHECK_TIMINGS_` defaults to a value from the environment even for `R CMD check --as-cran`; this allows for exceptionally fast or slow platforms.
  It now applies to checking PDF and HTML versions of the manuals, and 'checking CRAN incoming feasibility'.
- `R CMD check` can optionally (but included in '`--as-cran`') check whether HTML math rendering *via* KaTeX works for the package '`.Rd`' files.
- Non-interactive debugger invocations can be trapped by setting the environment variable `_R_CHECK_BROWSER_NONINTERACTIVE_` to a true value. This is enabled by `R CMD check --as-cran` to detect the use of leftover `browser()` statements in the package.
- The use of `sprintf` and `vsprintf` from C/C++ has been deprecated in macOS 13 and is a known security risk. `R CMD check` now reports (on all platforms) if their use is found in compiled code: replace by `snprintf` or `vsnprintf` respectively. [**NB:** whether such calls get compiled into the package is platform-dependent.]
- Where recorded at installation, `R CMD check` reports the C and Fortran compilers used to build R.

It reports the OS in use (if known, as given by `osVersion`) as well as that R was built for.
It notes if a C++ standard was specified which is older than the current default: many packages have used C++11 to mean 'not C++98' — as C++11 is the minimum supported since R 4.0.0, that specification can probably be removed.

- `R CMD INSTALL` reports the compilers (and on macOS, the SDK) used, and this is copied to the output of `R CMD check`.
  Where a C++ standard is specified, it is reported.
- `R CMD check`'s 'checking compilation flags in Makevars' has been relaxed to accept the use of flags such as '`-std=f2008`' in '`PKG_FFLAGS`'.
- `tools::buildVignettes()` has a new argument `skip`, which is used by `R CMD check` to skip (and note) vignettes with unavailable '`\VignetteDepends`' (PR#18318).
- New generic `.AtNames()` added to enable class-specific completions after `@`. The formerly internal function `findMatches()` is now exported, mainly for use in methods for `.DollarNames()` and `.AtNames()`.

DEPRECATED AND DEFUNCT:

- `default.stringsAsFactors()` is defunct.
- Calling `as.data.frame.<class>()` directly (for 12 atomic classes) is going to be formally deprecated, currently activated by setting the environment variable `_R_CHECK_AS_DATA_FRAME_EXPLICIT_METHOD_` to non-empty, which also happens in `R CMD check --as-cran`.

BUG FIXES:

- Hashed `environment`s with sizes less than 5 can now grow. (Reported to R-devel by Duncan Garmonsway.)
- `as.character(<Rd>, deparse = TRUE)` failed to re-escape curly braces in LaTeX-like text. (Reported by Hadley Wickham in PR#18324.)
- `library()` now passes its `lib.loc` argument when requiring `Depends` packages; reported (with fix) in PR#18331 by Mikael Jagan.
- `R CMD Stangle`: improved message about 'Output' files.
- `head(x, n)` and `tail(x, n)` now signal an error if `n` is not numeric, instead of incidentally "working" sometimes returning all of `x`. Reported and discussed by Colin Fay, in PR#18357.
- The `"lm"` method for `summary()` now gives the correct F-statistic when the model contains an offset. Reported in PR#18008.
- `C()` and `` `contrasts<-`() `` now preserve factor level names when given a function object (as opposed a function name which did preserve names). Reported in PR#17616.
- `c(a = 1, 2)[[]]` no longer matches `2` but rather signals a *classed* error. Reported and analysed by Davis Vaughan in PR#18367, a duplicate of PR#18004, by Jan Meis et al. For consistency, `NULL[[]]` is also erroneous now. `x[[]] <- v` gives an error of the same class `"MissingSubscriptError"`.
- The `relist()` function of **utils** now supports `NULL` elements in the skeleton (PR#15854).
- `ordered(levels = *)` (missing `x`) now works analogously to `factor(, ordered=TRUE)`; reported (with fix) by Achim Zeileis in PR#18389.
- User-defined Rd macro definitions can now span multiple lines, thanks to a patch from Duncan Murdoch. Previously, the Rd parser silently ignored everything after the first line.

- Plain-text help (`tools::Rd2txt()`) now preserves an initial blank line for text following description list items.
- `tools::Rd2HTML()` and `tools::Rd2latex()` no longer split `\arguments` and `\value` lists at Rd comments.
- `tools::Rd2latex()` now correctly handles optional text outside `\item`s of argument lists as well as bracketed text at the beginning of sections, e.g., `\value{[NULL]}`.
- `as.character(<POSIXt>)` now behaves more in line with the methods for atomic vectors such as numbers, and is no longer influenced by `options()`. Ditto for `as.character(<Date>)`. The `as.character()` method gets arguments `digits` and `OutDec` with defaults *not* depending on `options()`. Use of `as.character(*, format = .)` now warns.
- Similarly, the `as.character.hexmode()` and `*.octmode()` methods also behave as `good citizen` methods and back compatibility option `keepStr = TRUE`.
- The `as.POSIXlt(<POSIXlt>)` and `as.POSIXct(<POSIXct>)` default methods now do obey their `tz` argument, also in this case.
- `as.POSIXlt(<Date>)` now does apply a `tz` (time zone) argument, as does `as.POSIXct()`; partly suggested by Roland Fuß on the R-devel mailing list.
- `as.Date.POSIXlt(x)` now also works when the list components are of unequal length, aka "partially filled" or "ragged".
- `expand.model.frame()` looked up variables in the wrong environment when applied to models fitted without `data`. Reported in PR#18414.
- `time()` now (also) uses the `ts.eps = getOption("ts.eps")` argument and thus by default rounds values very close to the start (or end) of a year. Based on a proposal by Andreï V. Kostyrka on R-help.
- Printing of a `factanal()` result with just one factor and `sort = TRUE` now works regularly, fixing PR#17863 by Timothy Bates, thanks to the 'R Contributors' working group.
- Printing 0-length objects of class `"factor"`, `"roman"`, `"hexmode"`, `"octmode"`, `"person"`, `"bibentry"`, or `"citation"` now prints something better, one of which fixes PR#18422, reported by Benjamin Feakins.
- `Sys.timezone()` queries `timedatectl` only if `systemd` is loaded; addressing a report by Jan Gorecki in PR#17421.
- The formula method of `cor.test()` had scoping problems when `environment(formula)` was not the calling environment; reported with a patch proposal by Mao Kobayashi in PR#18439.
- `attach()` of an environment with active bindings now preserves the active bindings. Reported by Kevin Ushey in PR#18425.
- BLAS detection now works also with system-provided libraries not available as regular files. This fixes detection of the Accelerate framework on macOS since Big Sur. Reported by David Novgorodsky.
- `download.file()` gives a helpful error message in case of an invalid `download.file.method` option, thanks to Colin Fay's report in PR#18455.
- Sporadic crashes of `Rterm` when using completion have been fixed.
- `Rprof()` is now more reliable. A livelock in thread initialization with too short sampling interval has been fixed on macOS. A deadlock in using the C runtime has been fixed on Windows. A potential deadlock has been prevented on Unix.
- Cursor placement in `Rgui` now works even after a fixed-width font is selected.
- Mandatory options (`options()`) are now set on startup so that saving and restoring them always works (PR#18372).

- Package installation, `R CMD INSTALL` or `install.packages(*)`, now parses each of the '`<pkg>/R/*.R`' files individually instead of first concatenating and then `parse()`ing the large resulting file. This allows parser or syntax errors to be diagnosed with correct file names and line numbers, thanks to Simon Dedman's report and Bill Dunlap's patch in PR#17859.
  This *does* require syntactically self contained R source files now, fixing another inadvertent bug.
- `predict.lm(<model with offset>)` now finds the offset in the correct environment, thanks to André Gillibert's report and patch in PR#18456.
- `getInitial(<formula>)` now finds the `selfStart` model in the correct environment. (Reported by Ivan Krylov in PR#18368.)
- Fix for possible segfault when using recently-added graphics features, such as gradients, clipping paths, masks, and groups with `pdf(file=NULL)`.
- `class(m) <- class(m)` no longer changes a matrix `m` by adding a class *attribute*.
- `packageDate(pkg)` now only warns once if there is no `pkg`.
- When `ts()` creates a multivariate time series, `"mts"`, it also inherits from `"array"` now, and `is.mts()` is documented *and* stricter.
- `Rd2txt()` now preserves line breaks of `\verb` Rd content and from duplicated `\cr`. The former also fixes the rendering of verbatim output from Rd `\Sexpr` in plain-text help.
- `uniroot(f, interval)` should no longer wrongly converge *outside* the interval in some cases where `abs(f(x)) == Inf` for an `x` at the interval boundary, thanks to posts by Ben Bolker and Serguei Sokol on R-devel.
- Vectorized alpha handling in palette functions such as in `gray()`, `rainbow()`, or `hcl.colors()` works correctly now, thanks to Achim Zeileis' report and patch in PR#18476.
- Formatting and `print()`ing of `bibentry` objects has dropped the deprecated `citation.bibtex.max` argument, such that the `bibtex` argument's default for `print.bibentry()` depends directly on the `citation.bibtex.max` option, whereas in `format.bibentry()` the option no longer applies.
- Attempting to use a character string naming a foreign function entry point in a foreign function call in a package will now signal an error if the packages has called `R_forceSymbols` to specify that symbols must be used.
- An error in `table()` could permanently set `options(warn=2)` promoting all subsequent warnings to errors.
- The `sigma()` function gave misleading results for binary GLMs. A new method for objects of class `"glm"` returns the square root of the estimate of the dispersion parameter using the same calculation as `summary.glm()`.
- `bs()` and `ns()` in the (typical) case of automatic knot construction, when some of the supposedly inner knots coincide with boundary knots, now moves them inside (with a warning), building on PR#18442 by Ben Bolker.
- `R CMD` on Windows now skips the site profile with '`--no-site-file`' and '`--vanilla`' even when `R_PROFILE` is set (PR#18512, from Kevin Ushey).



CHANGES IN R 4.2.3

C-LEVEL FACILITIES:

- The definition of `DL_FUNC` in '`R_ext/Rdynload.h`' has been changed to be fully C-compliant. This means that functions loaded *via* for example `R_GetCCallable` need to be cast to an appropriate type if they have any arguments.

- `.Machine` has a new element `sizeof.time_t` to identify old systems with a 32-bit type and hence a limited range of date-times (and limited support for dates millions of years from present).

PACKAGE INSTALLATION:

- (Windows) The default C++ standard had accidentally been left at C++11 when it was changed to C++14 on Unix.

BUG FIXES:

- As `"POSIXlt"` objects may be "partially filled" and their list components meant to be recycled, `length()` now is the length of the longest component.
- `as.POSIXlt.Date()` could underflow for dates in the far past (more than half a million years BCE).
- `as.Date.POSIXlt(x)` would return `"1970-01-01"` instead of `NA` in R 4.2.2, e.g., for

  ```
  x <- as.POSIXlt(c("2019-01-30","2001-1-1"))
  x$mon <- c(0L, NA); as.Date(x)
  ```
- `R CMD check` failed to apply enabled `_R_CHECK_SUGGESTS_ONLY_` to examples and vignettes (regression in R 4.2.0).
- `R CMD check` did not re-build vignettes in separate processes by default (regression in R 4.2.0).
- Running examples from HTML documentation now restores previous **knitr** settings and options (PR#18420).
- Quartz: fonts are now located using Core Graphics API instead of deprecated ATS which is no longer supported in the macOS 13 SDK (PR#18426). This also addresses an issue where the currently used font in the Quartz device context was not correctly retained.
- (Windows) Math symbols in text drawing functions are again rendered correctly (PR#18440). This fixes a regression in R 4.2.1 caused by a fix in PR#18382 which uncovered an issue in GraphApp due to which the symbol charset was not used with TT Symbol font face.
- (Windows) Installing a package with a '`src/Makefile.{win,ucrt}`' file includes '`~/.R/Makevars.win64`' in the search for user makevars, as documented in "R Installation and Administration" and done for packages with a '`src/Makevars.{win,ucrt}`' file.
- `format(<POSIXlt_w/_unbalanced_sec>, "....%OS<n>")` with $n > 0$ no longer accidentally uses the unbalanced seconds, thanks to Suharto Anggono's report (including patch) in PR#18448.
- `solve.default(a, b)` works around issues with some versions of LAPACK when `a` contains `NA` or `NaN` values.
- When `UseMethod()` cannot dispatch, it no longer segfaults producing the error message in case of a long `class()`, thanks to Joris Vankerschaver's report (including patch) in PR#18447.
- When `example(foo, ..)` produces graphics on an interactive device it needs to open itself, it now leaves `devAskNewPage()` unchanged even when it was `FALSE`, thus fixing a 14 years old '<FIXME>'.
- `packageDescription()` again catches errors from encoding conversions. This also fixes broken `packageVersion()` in C locale on systems where `iconv` does not support transliteration.



CHANGES IN R 4.2.2

NEW FEATURES:

- `tools::Rdiff(useDiff = TRUE)` checks for the presence of an external `diff` command and switches to `useDiff = FALSE` if none is found. This allows `R CMD Rdiff` to always work.
- On Windows, environment variable `R_LIBCURL_SSL_REVOKE_BEST_EFFORT` can be used to switch to only 'best-effort' SSL certificate revocation checks with the default `"libcurl"` download method. This reduces security, but may be needed for downloads to work with MITM proxies (PR#18379).
- (macOS) The run-time check for libraries from XQuartz for X11 and Tcl/Tk no longer uses `otool` from the Apple Developer Tools (PR#18400).
- The LaTeX style for producing the PDF manuals, '`Rd.sty`', now loads the standard '`amsmath`', '`amsfonts`' and '`amssymb`' packages for greater coverage of math commands in the Rd `\eqn` and `\deqn` macros. The `\mathscr` LaTeX command is also provided (via the '`mathrsfs`' package, if available, or the '`amsfonts`' bundle otherwise), fulfilling the wish of PR#18398.
- (Windows) The default format of `readClipboard()` and `writeClipboard()` has been changed to `13` (`CF_UNICODETEXT`).

INSTALLATION on a UNIX-ALIKE:

- The PDF manuals (if built) can be compacted by the new target `make compact-pdf` (at the top level or in directory '`doc/manual`').
- There is now `configure` support for LLVM clang 15 on Linux, which defaults to position-independent (PIE) executables whereas `gfortran` does not.
- Many small changes to ease compilation (and suppress warnings) with LLVM `clang` 15.

BUG FIXES:

- `Rscript -e` would fail if '`stdin`' were closed (Reported by Henrik Bengtsson.)
- `qt(*, log.p=TRUE)` in outer tails no longer produces `NaN` in its final steps, thus fixing PR#18360.
- `tools::Rd2latex()` now escapes hashes and ampersands when writing URLs, fixing LaTeX errors with such URLs in `\tabular`.
- When `isGeneric(f, fdef=*)` is used with mismatching names, the warning is better understandable; reported (with fix) in PR#18370 by Gabe Becker.
- `poly(x, n)` now works again (and is now documented) when `x` is a `"Date"` or `"POSIXct"` object, or of another class while fulfilling `mode(x) == "numeric"`. This also enables `poly(x, *, raw=TRUE)` for such variables. Reported by Michael Chirico to R-devel.
- `write.table()`, `write.csv()` and `write.csv2()` restore their numerical precision (internal equivalent of `digits = 15`) after an interrupt (PR#18384).
- One can now read also byte `FF` from a clipboard connection (PR#18385).
- `source("")` and `source(character())` now give more helpful error messages.
- `R CMD check --as-cran` set `_R_CHECK_TIMINGS_` too late to have the intended effect.
- `as.POSIXlt(x)` now also works with very large dates `x`, fixing PR#18401 reported by Hannes Mühleisen.
- Files can now be extracted even from very large zip archives (PR#18390, thanks to Martin Jakt).

- Non-finite objects of class `"POSIXlt"` are now correctly coerced to classes `"Date"` and `"POSIXct"`; following up on the extension to `format()` them correctly.
- Added methods for `is.finite()`, `is.infinite()` and `is.nan()` for `"POSIXlt"` date-time objects.

BUG FIXES on Windows:

- Non-ASCII characters are now properly displayed on Windows in windows created using GraphApp via e.g. `winDialogString` thanks to a workaround for an at least surprising Windows behavior with UTF-8 as the system encoding (PR#18382).
- Find and replace operations work again in the script editor in `Rgui` on Windows.
- Computation of window size based on requested client size in GraphApp when running in a multi-byte locale on Windows has been fixed (regression in R 4.2.0 for users of systems where R 4.1 used a single-byte locale). `Rgui` again respects the number of console rows and columns given in '`Rconsole`' file.
- `Rterm` support for `Alt+xxx` sequences has been fixed to produce the corresponding character (only) once. This fixes pasting text with tilde on Italian keyboard (PR#18391).



CHANGES IN R 4.2.1

NEW FEATURES:

- New function `utils::findCRANmirror()` to find out if a CRAN mirror has been selected, otherwise fallback to the main site. This behaves in the same way as `tools::CRAN_package_db()` and is intended for packages wishing to access CRAN for purposes other than installing packages.
  The need for this was shown by a day when the main CRAN website was offline and a dozen or so packages which had its URL hardcoded failed their checks.

INSTALLATION on a UNIX-ALIKE:

- The libraries searched for by '`--with-blas`' (without a value) now include BLIS (after OpenBLAS but before ATLAS). And on macOS, the Accelerate framework (after ATLAS). (This is patterned after the `AX_BLAS` macro from the Autoconf Archive.)
- The included LAPACK sources have been updated to 3.10.1.

UTILITIES:

- The (full path to) the command `tidy` to be used for HTML validation can be set by environment variable `R_TIDYCMD`.
- Setting environment variable `_R_CHECK_RD_VALIDATE_RD2HTML_` to a false value will override `R CMD check --as-cran` and turn off HTML validation. This provides a way to circumvent a problematic `tidy`.
  The 2006 version that ships with macOS is always skipped.

C-LEVEL FACILITIES:

- The undocumented legacy declarations of `Sint`, `Sfloat`, `SINT_MAX` and `SINT_MIN` in header '`R.h`' are deprecated.

BUG FIXES:

- `fisher.test(d)` no longer segfaults for "large" `d`; fixing PR#18336 by preventing/detecting an integer overflow reliably.

- `tar(., files=*)` now produces correctly the warning about invalid uid or gid of files, fixing PR#18344, reported by Martin Morgan.
- `tk_choose.files()` with `multi = FALSE` misbehaved on paths containing spaces (PR#18334) (regression introduced in R 4.0.0).
- `sort(x, partial = ind, *)` now works correctly notably for the non-default `na.last = FALSE` or `TRUE`, fixing PR#18335 reported by James Edwards.
- Environment variable `_R_CHECK_XREFS_REPOSITORIES_` is only used for checking '`.Rd`' cross-references in `R CMD check` (as documented) and not for other uses looking for a CRAN mirror.
- The search for a CRAN mirror when checking packages now uses `getOption("repos")` if that specifies a CRAN mirror, even when it does not also specify all three Bioconductor repositories (as was previously required).
- The HTML code generated by `tools::Rd2HTML()` has been improved to pass `tidy` 5.8.0.

BUG FIXES on Windows:

- Writing to a clipboard connection works again, fixing a regression in R 4.2.0 (PR#18332). Re-using a closed clipboard connection longer issues a spurious warning about an ignored encoding argument.
- C function `getlocale` no longer attempts to query an unsupported category from the OS, even when requested at R level, which may cause crashes when R 4.2.0 (which uses UCRT) is embedded (reported by Kevin Ushey).
- Accent keys now work in GraphApp Unicode windows, which are used by `Rgui` whenever running in a multibyte locale (so also in UTF-8, hence fixing a regression in R 4.2.0 for users of systems where R 4.1 used a single-byte locale).
- Completion in `Rgui` now works also with non-ASCII characters.
- `Rgui` no longer truncates usage information with '`--help`'.
- Text injection from external applications via `SendInput` now works in GraphApp Unicode windows, fixing a regression in R 4.2.0 for `Rgui` users of systems where R 4.1 used a single-byte locale but R 4.2.0 uses UTF-8.
- Performance of `txtProgressBar()` in `Rgui` when running in a multi-byte locale has been improved (fixing a performance regression in R 4.2.0 for users of systems where R 4.1 used a single-byte locale).
- The script editor in `Rgui` now works also on systems using UTF-8 as the native encoding. Users of the script editor have to convert their scripts with non-ASCII characters to UTF-8 before reading them in R 4.2.1 or newer (on recent Windows where UTF-8 is used). This fixes a regression in R 4.2.0, which prevented some operations with scripts when they contained non-ASCII characters.



CHANGES IN R 4.2.0

SIGNIFICANT USER-VISIBLE CHANGES:

- The `formula` method of `aggregate()` now matches the generic in naming its first argument `x` (resolving PR#18299 by Thomas Soeiro).
  This means that calling `aggregate()` with a formula as a named first argument requires name `formula` in earlier versions of R and name `x` now, so portable code should not name the argument (code in many packages did).
- Calling `&&` or `||` with either argument of length greater than one now gives a warning (which it is intended will become an error).

- Calling `if()` or `while()` with a condition of length greater than one gives an error rather than a warning. Consequently, environment variable `_R_CHECK_LENGTH_1_CONDITION_` no longer has any effect.
- Windows users should consult the WINDOWS section below for some profound changes including
  - Support for 32-bit builds has been dropped.
  - UTF-8 locales are used where available.
  - The default locations for the R installation and personal library folder have been changed.

  Thanks to Tomas Kalibera for months of work on the Windows port for this release.

NEW FEATURES:

- `matrix(x, n, m)` now warns in more cases where `length(x)` differs from `n * m`, as suggested by Abby Spurdle and Wolfgang Huber in Feb 2021 on the R-devel mailing list.
  This warning can be turned into an error by setting environment variable `_R_CHECK_MATRIX_DATA_` to 'TRUE': `R CMD check --as-cran` does so unless it is already set.
- Function `file_test()` in package **utils** gains tests for symlinks, readability and writability.
- `capabilities("libxml")` is now false.
  The description of `capabilities("http/ftp")` now reflects that it refers to the default method, no longer the internal one.
- `simplify2array()` gains an `except` argument for controlling the exceptions used by `sapply()`.
- Environment variables `R_LIBS_USER` and `R_LIBS_SITE` are both now set to the R system default if unset or empty, and can be set to `NULL` to indicate an empty list of user or site library directories.
- The warning for `axis()`(-like) calls in cases of relatively small ranges (typically in log-scale situations) is slightly improved *and* suppressed from explicit calls to `.axisPars()` as has always been the intention.
- The `contrasts` setter function `` `contrasts<-` `` gains an explicit default `how.many = NULL` rather than just using `missing(how.many)`.
- `grid.pretty()` gains a new optional argument `n = 5`.
- There is a new function `.pretty()` with option `bounds` as a technical-utility version of `pretty()`. It and `pretty()` gain a new argument `f.min` with a better than back-compatible default.
- Function `grDevices::axisTicks()` and related functions such as `graphics::axis()` work better, notably for the log scale; partly because of the `pretty()` improvements, but also because care is taken e.g., when `ylim` is finite but `diff(ylim)` is infinite.
- `nclass.FD()` gains a `digits` option.
- The R Mathlib internal C function `bd0()` (called indirectly from a dozen probability density and distribution functions such as `dpois()`, `dbinom()`, `dgamma()`, `pgamma()` *etc*) has been complemented by a more sophisticated and (mostly) more accurate C function `ebd0()`, currently called only by internal `dpois_raw()` improving accuracy for R level `dpois()` and potentially others calling it such as `dnbinom()`, `dgamma()` or `pgamma()`. (Thanks to Morten Welinder's PR#15628.)
- `write.ftable()` gains `sep = " "` argument as suggested by Thomas Soeiro.
- The names of the locale categories supported by R's `Sys.getlocale()` and `Sys.setlocale()` are now provided by variable `.LC.categories` in the `base` namespace.

- The `Date` and `POSIXt` methods for `hist()` and the `histogram` method for `plot()` now also use the new default `col = "lightgray"` in consistency with the corresponding change to `hist()`'s default for R 4.0.0.
- `hist.default()` gains new `fuzz` argument, and the histogram `plot` method no longer uses fractional axis ticks when displaying counts (`"Frequency"`).
- `mapply()` and hence `Map()` now also obey the "max-or-0-if-any" recycling rule, such that, e.g., `Map(`+`, 1:3, 1[0])` is valid now.
- `as.character(<obj>)` for `"hexmode"` or `"octmode"` objects now fulfils the important basic rule `as.character(x)[j] === as.character(x[j])`.
- The set utility functions, notably `intersect()` have been tweaked to be more consistent and symmetric in their two set arguments, also preserving a common `mode`.
- `substr(ch, start,end) <- new` now e.g., preserves `names(ch)`; ditto for `substring()`, thanks to a patch from Brodie Gaslam.
- `plot(<lm>)` gains a `extend.ylim.f` argument, in partial response to PR#15285; further PR#17784 is fixed thanks to several contributors and a patch by Elin Waring. The Cook's dist contours get customizable via `cook.col` and `cook.lty` with a different default color and their legend is nicer by default and customizable via `cook.legendChanges`.
- Attempting to subset an object that is not subsettable now signals an error of class `notSubsettableError`. The non-subsettable object is contained in the `object` field of the error condition.
- Subscript-out-of-bounds errors are now signaled as errors of class `subscriptOutOfBoundsError`.
- Stack-overflow errors are now signaled as errors inheriting from class `stackOverflowError`. See `?stackOverflowError` for more details.
- New partly experimental `Sys.setLanguage()` utility, solving the main problem of PR#18055.
- `gettext()` and `gettextf()` get a new option `trim = TRUE` which when set to false allows translations for strings such as `"Execution halted\n"` typical for C code.
- An experimental implementation of hash tables is now available. See `?hashtab` for more details.
- `identical()` gains a `extptr.as.ref` argument for requesting that external pointer objects be compared as reference objects.
- `reorder()` gets an argument `decreasing` which it passes to `sort()` for level creation; based on the wish and patch by Thomas Soeiro in PR#18243.
- `as.vector()` gains a `data.frame` method which returns a simple named list, also clearing a long standing 'FIXME' to enable `as.vector(<data.frame>, mode="list")`. This breaks code relying on `as.vector(<data.frame>)` to return the unchanged data frame.
- `legend()` is now vectorized for arguments `cex`, `x.intersp`, and `text.width`. The latter can now also be specified as a vector (one element for each column of the legend) or as `NA` for computing a proper column wise maximum value of `strwidth(legend)`. The argument `y.intersp` can be specified as a vector with one entry for each row of the legend.
  `legend()` also gains new arguments `title.cex` and `title.font`. Thanks to Swetlana Herbrandt.
- Deparsing no longer remaps attribute names `dim`, `dimnames`, `levels`, `names` and `tsp` to historical S-compatible names (which `structure()` maps back).
- `sample()` and `sample.int()` have additional sanity checks on their `size` and `n` arguments.

`all.equal.numeric()` gains a sanity check on its `tolerance` argument – calling `all.equal(a, b, c)` for three numeric vectors is a surprisingly common error.
`mean(na.rm =)`, `rank(na.last =)`, `barplot(legend.text =)`, `boxplot()`, `contour(drawlabels =)`, `polygon(border =)` and `methods::is(class2 =)` have more robust sanity checks on their arguments.
`R CMD Rd2pdf` (used by `R CMD check`) has a more robust sanity check on the format of `\alias{}` commands.

- `psigamma(x, deriv)` for negative `x` now also works for `deriv = 4` and `5`; their underlying C level `dpsifn()` is documented in 'Writing R Extensions'.
- The HTML help system now uses HTML5 (wish of PR#18149).
- `ks.test()` now provides exact p-values also with ties and MC p-values in the two-sample (Smirnov) case. By Torsten Hothorn.
- `ks.test()` gains a formula interface, with `y ~ 1` for the one-sample (Kolmogorov) test and `y ~ group` for the two-sample (Smirnov) test. Contributed by Torsten Hothorn.
- The return value from `ks.test()` now has class `c("ks.test", "htest")` – packages using `try()` need to take care to use `inherits()` and not `==` on the class.
- New functions `psmirnov()`, `qsmirnov()` and `rsmirnov()` in package **stats** implementing the asymptotic and exact distributions of the two-sample Smirnov statistic.
- `iconv()` now allows `sub = "c99"` to use C99-style escapes for UTF-8 inputs which cannot be converted to encoding `to`.
- In a forward pipe `|>` expression it is now possible to use a named argument with the placeholder `_` in the `rhs` call to specify where the `lhs` is to be inserted. The placeholder can only appear once on the `rhs`.
- The included LAPACK sources have been updated to version 3.10.0, except for the four Fortran 77 routines which 3.10.0 has re-implemented in Fortran 90 (where the older versions have been retained as the `R` build process does not support Fortran 90).
- `path.expand()` and most other uses of tilde expansion now warn if a path would be too long if expanded. (An exception is `file.exists()`, which silently returns false.)
- `trunc(<Date>, *)` now supports `units = "months"` or `"years"` for consistency with the `POSIXt` method, thanks to Dirk Eddelbuettel's proposal in PR#18099.
- `list2DF()` now checks that its arguments are of the same length, rather than use recycling.
- The HTML help system has several new features: LaTeX-like math can be typeset using either KaTeX or MathJax, usage and example code is highlighted using Prism, and for dynamic help the output of examples and demos can be shown within the browser if the **knitr** package is installed. These features can be disabled by setting the environment variable `_R_HELP_ENABLE_ENHANCED_HTML_` to a false value.

GRAPHICS:

- The graphics engine version, `R_GE_version`, has been bumped to `15` and so packages that provide graphics devices should be reinstalled.
- The **grid** package now allows the user to specify a "vector" of pattern fills. The `fill` argument to `gpar()` accepts a list of gradients and/or patterns and the functions `linearGradient()`, `radialGradient()`, and `pattern()` have a new `group` argument. Points grobs (data symbols) can now also have a pattern fill.
  The `grobCoords()` function now returns a more informative and complex result.
- The **grid** package has new functions for drawing isolated groups: `grid.group()`, `grid.define()`, and `grid.use()`. These functions add compositing operators and affine transformations to R's graphics capabilities.

The **grid** package also has new functions for stroking and filling paths: `grid.stroke()`, `grid.fill()`, and `grid.fillStroke()`.
A new function `as.path()` allows the user to specify the fill rule for a path that is to be used for clipping, stroking, or filling; available options are `"winding"` and `"evenodd"`. A new function `as.mask()` allows the user to specify the type of a mask; available options are `"alpha"` and `"luminance"`.
These new features are only supported so far (at most) on the Cairo-based graphics devices and on the `pdf()` device.

- `dev.capabilities()` reports on device support for the new features.
- `par()` now warns about unnamed non-character arguments to prevent misuse such as `{usr <- par("usr"); par(usr)}`.

WINDOWS:

- R uses UTF-8 as the native encoding on recent Windows systems (at least Windows 10 version 1903, Windows Server 2022 or Windows Server 1903). As a part of this change, R uses UCRT as the C runtime. UCRT should be installed manually on systems older than Windows 10 or Windows Server 2016 before installing R.
- The default personal library on Windows, folder '`R\win-library\x.y`' where '`x.y`' stands for R release '`x.y.z`', is now a subdirectory of Local Application Data directory (usually a hidden directory '`C:\Users\username\AppData\Local`'). Use `shell.exec(.libPaths()[1])` from R to open the personal library in Explorer when it is first in the list (PR#17842).
- R uses a new 64-bit Tcl/Tk bundle. The previous 32-bit/64-bit bundle had a different layout and can no longer be used.
- Make files and installer scripts for Windows have been tailored to '`Rtools42`', the newly recommended 64-bit `gcc` 10.3 MinGW-W64 UCRT toolchain.
- '`Rtools42`' by default uses the Windows security features ASLR and DEP; hence CRAN builds of R and packages also do.
- R now supports files '`Makevars.ucrt`', '`Makefile.ucrt`', '`configure.ucrt`' and '`cleanup.ucrt`' in packages, which are used in preference to the '`.win`' variants. This allows keeping the '`.win`' files around to support older versions of R. This feature will be removed in the future once support for older versions of R would no longer be needed.
- `R.version` gains a new field `crt` (only on Windows) to denote the C runtime. The value is `"ucrt"`.
- On Windows, `download.file(method = "auto")` and `url(method = "default")` now follow Unix in using `"libcurl"` for all except '`file://`' URIs.
- '`Rtools42`' includes an unpatched Msys2 build of GNU `tar`. Paths including drive letters can be made to work by adding '`--force-local`' to environment variable `TAR_OPTIONS`. ('`Rtools40`' and earlier included a patched version which defaulted to this option.)
- Installer builds of R automatically find the '`Rtools42`' software collection as well as the compiler toolchain. No `PATH` setting is required from the user.
- The default installation directory of R for a user-only installation has been changed to the User Program Files directory (usually a hidden directory '`C:\Users\username\AppData\Local\Programs`') to follow Windows conventions. Use `shell.exec(R.home())` from R to open the R installation directory in Explorer (PR#17842).
- R now supports installation-time patching of packages. Patches may be installed from a supplied URL or a local directory or disabled. Patches are included into the

installed packages for reference. This experimental feature may be removed in the future.

- `libcurl` is now required for building from source.
- The clipboard connection now works also with text in other than the current native encoding (PR#18267, with Hiroaki Yutani). Text is always pasted to the clipboard in UTF16-LE and the `encoding` argument is ignored.
- The internal case-changing functions are now used by default on Windows – this circumvents problems (for example with E acute) of the UCRT Windows' runtime.
- R on Windows now uses the system memory allocator. Doug Lea's allocator was used since R 1.2.0 to mitigate performance limitations seen with system allocators on earlier versions of Windows.
- `memory.limit()` and `memory.size()` are now stubs on Windows (as on Unix-alikes).
- Applications embedding R on Windows can now use additional callbacks, which have so far only been available only on Unix (PR#18286).

INSTALLATION:

- Facilities for accessing '`ftp://`' sites are no longer tested (except *pro tem* for `curlGetHeaders()`) as modern browsers have removed support.
- R can now be built with '`DEFS = -DSTRICT_R_HEADERS`' .

PACKAGE INSTALLATION:

- `R CMD INSTALL` no longer tangles vignettes. This completes an `R CMD build` change in R 3.0.0 and affects packages built before R 3.0.2. Such packages should be re-made with `R CMD build` to have the tangled R code of vignettes shipped with the tarball.
- `USE_FC_LEN_T` will become the default: this uses the correct prototypes for Fortran BLAS/LAPACK routines called from C/C++, and requires adjustment of most such calls – see 'Writing R Extensions' §6.6.1. (This has been supported since R 3.6.2.)
- Package installation speed for packages installed with `keep.source` has been improved. This resolve the issue reported by Ofek Shilon in PR#18236.

UTILITIES:

- `R CMD check` can optionally report files/directories left behind in home, '`/tmp`' (even though `TMPDIR` is set) and other directories. See the "R Internals" manual for details.
- `R CMD check` now reports byte-compilation errors during installation. These are not usually fatal but may result in parts of the package not being byte-compiled.
- `_R_CHECK_DEPENDS_ONLY_` can be applied selectively to examples, tests and/or vignettes in `R CMD check`: see the "R Internals" manual.
- `_R_CHECK_SRC_MINUS_W_IMPLICIT_` now defaults to true: recent versions of Apple `clang` on macOS have made implicit function declarations in C into a compilation error.
- `R CMD check --as-cran` makes use of the environment variable `AUTORECONF`. See the "R Internals" manual §8 for further details.
- `R CMD check --use-valgrind` also uses `valgrind` when re-building vignettes as some non-Sweave vignettes unhelpfully comment out all their code when `R CMD check` runs vignettes.
- Errors in re-building vignettes (unless there are LaTeX errors) are reported by `R CMD check` as '`ERROR`' rather than '`WARNING`' when running vignettes has been skipped (as it frequently is in CRAN checks and by '`--as-cran`').
- `R CMD Rd2pdf` gains a '`--quiet`' option that is used by `R CMD build` when building the PDF package manual.

- `R CMD Rd2pdf` now always runs LaTeX in batch mode, consistent with Texinfo $\geq$ 6.7. The '`--batch`' option is ignored.
- `R CMD build` and `R CMD check` now include the Rd file name and line numbers in the error message of an `\Sexpr` evaluation failure.
- For packages using the `\doi` Rd macro (now an install-time `\Sexpr`) but no other dynamic Rd content, `R CMD build` now produces a smaller tarball and is considerably faster – skipping temporary package installation.
- `R CMD check` can optionally (but included in '`--as-cran`') validate the HTML produced from the packages '`.Rd`' files. See https://blog.r-project.org/2022/04/08/enhancements-to-html-documentation/: this needs a fairly recent version of HTML Tidy to be available.

C-LEVEL FACILITIES:

- The non-API header '`R_ext/R-ftp-http.h`' is no longer provided, as the entry points it covered are now all defunct.
- A number of non-API declarations and macro definitions have been moved from the installed header '`Rinternals.h`' to the internal header '`Defn.h`'. Packages that only use entry points and definitions documented to be part of the API as specified in 'Writing R Extensions' §6 should not be affected.
- The macro `USE_RINTERNALS` no longer has any effect when compiling package code. Packages which also use `R_NO_REMAP` will need to ensure that the remapped names are used for calls to API functions that were formerly also made available as macros.
- The deprecated legacy S-compatibility macros `PROBLEM`, `MESSAGE`, `ERROR`, `WARN`, `WARNING`, `RECOVER`, ... are no longer defined in '`R_exts/RS.h`' (included by '`R.h`'). Replace these by calls to `Rf_error` and `Rf_warning` (defined in header '`R_ext/Error.h`' included by '`R.h`').
  Header '`R_ext/RS.h`' no longer includes '`R_ext/Error.h`'.
- Header '`R_ext/Constants.h`' (included by '`R.h`') when included from C++ now includes the C++ header '`cfloat`' rather than the C header '`float.h`' (now possible as C++11 is required).
- The legacy S-compatibility macros `DOUBLE_*` in '`R_ext/Constants.h`' (included by '`R.h`') are deprecated.
- The deprecated S-compatibility macros `SINGLE_*` in '`R_ext/Constants.h`' (included by '`R.h`') have been removed.
- `R_Calloc`, `R_Free` and `R_Realloc` are preferred to their unprefixed forms and error messages now use the prefix. These forms were introduced in R 3.4.0 and are available even when `STRICT_R_HEADERS` is defined.
- `rmultinom` has been documented in 'Writing R Extensions' §6 so is now part of the R API.
- Similarly, `Rtanpi`, called from R level `tanpi()` is now part of the R API.
- The long-deprecated, undocumented and non-API entry point `call_R` is no longer declared in '`R_ext/RS.h`' (included by '`R.h`').
- The header '`S.h`' which has been unsupported since Jan 2016 has been removed. Use '`R.h`' instead.

DEPRECATED AND DEFUNCT:

- The (non-default and deprecated) `method = "internal"` for `download.file()` and `url()` no longer supports '`http://`' nor '`ftp://`' URIs. (It is used only for '`file://`' URIs.)
  On Windows, `download.file(method = "wininet")` no longer supports '`ftp://`' URIs. (It is no longer the default method, which is `"libcurl"` and does.)

On Windows, the deprecated `method = "wininet"` now gives a warning for '`http://`' and '`https://`' URIs for both `download.file()` and `url()`. (It is no longer the default method.)

- On Windows, the command-line option '`--max-mem-size`' and environment variable `R_MAX_MEM_SIZE` are defunct. The memory allocation limit was important for 32-bit builds, but these are no longer supported.
- `default.stringsAsFactors()` is now formally deprecated, where that was only mentioned on its regular help page, previously. So it now gives a warning if called.
- `unix.time()` is defunct now; it had been deprecated since R 3.4.0.

BUG FIXES:

- Setting `digits = 0` in `format()`, `print.default()` (and hence typically `print()`) or `options()` is again invalid. Its behaviour was platform-dependent, and it is unclear what "zero significant digits" should mean (PR#18098).
- Messages from C code in the '`cairo`' section of package **grDevices** are now also offered for translation, thanks to Michael Chirico's PR#18123.
- `mean(x)` with finite `x` now is finite also without "long.double" capability.
- `R CMD Rd2pdf` no longer leaves an empty build directory behind when it aborts due to an already existing output file. (Thanks to Sebastian Meyer's PR#18141.)
- `density(x, weights = w, na.rm = TRUE)` when `anyNA(x)` is true, now removes weights "in parallel" to `x`, fixing PR#18151, reported by Matthias Gondan. Additionally, it gets a `subdensity` option.
- Conversion of `\Sexpr[]{<expR>}` to LaTeX or HTML no longer produces long blocks of empty lines when `<expR>` itself contains several lines all producing empty output. Thanks to a report and patch by Ivan Krylov posted to R-devel.
- `R CMD build` no longer fails if a package vignette uses child documents and '`inst/doc`' exists. (Thanks to Sebastian Meyer's PR#18156.)
- When an R documentation ('help' source) file '`man/foo.Rd`' in a package has `\donttest{..}` examples with a syntax error, it is now signalled as ERROR and with correct line numbers relating to the '`*-Ex.R`' file, thanks to Duncan Murdoch and Sebastian Meyer's reports and patch proposals in PR#17501.
- Improved determination of the correct translation domain in non-base packages, addressing the combination of PR#18092 and PR#17998 (`#c6`) with reports and *augmented* patch #2904 by Suharto Anggono.
  Note that `"R-base"` is no longer the default domain e.g., for top-level calls to `gettext()`; rather translation needs explicit `domain = *` specification in such cases.
- `identical(attrib.as.set=FALSE)` now works correctly with data frames with default row names (Thanks to Charlie Gao's PR#18179).
- `txtProgressBar()` now enforces a non-zero width for argument `char`, without which no progress can be visible.
- `dimnames(table(d))` is more consistent in the case where `d` is a list with a single component, thanks to Thomas Soeiro's report to R-devel.
  Further, `table(d1, d2)` now gives an error when `d1` and `d2` are data frames as suggested by Thomas in PR#18224.
- Fix for drawing semi-transparent lines and fills on the native Windows graphics device (PR#18219 and PR#16694). Thanks to Nick Ray for helpful diagnosis on Bugzilla.
- The deparser now wraps sub-expressions such as `if(A) ..` with parentheses when needed; thanks to Duncan Murdoch's PR#18232 and Lionel Henry's patches there.
- `remove.packages()` no longer tries to uninstall `Priority: base` packages, thanks to a report and suggestions by Colin Fay in PR#18227.

- `win.metafile()` now has `xpinch` and `ypinch` arguments so that the user can override Windows' (potentially wrong) guess at device dimensions.
- `x[i]` and `x[[i]]` for non-integer `i` should now behave in all cases as always documented: the index used is equivalent to `as.integer(i)` unless that would overflow where `trunc(i)` is used instead; thanks to Suharto Anggono's report and patch proposals in PR#17977.
- `asOneSidedFormula()` now associates the resulting formula with the global environment rather than the evaluation environment created for the call.
- `<bibentry>$name` now matches the field name case-insensitively, consistent with `bibentry()` creation and the replacement method.
- `cbind()` failed to detect some length mismatches with a mixture of time-series and non-time-series inputs.
- The default LaTeX style file '`Sweave.sty`' used by the `RweaveLatex` driver no longer loads the obsolete '`ae`' package; thanks to a report by Thomas Soeiro in PR#18271. Furthermore, it now skips '`\usepackage[T1]{fontenc}`' for engines other than pdfTeX (if detected) or if the new '`[nofontenc]`' option is used.
- `smooth.spline()` now stores its logical `cv` argument more safely, fixing a rare bug when printing, and also stores `n`.
- `smooth.spline(x,y,*)` now computes the `cv.crit` statistic correctly, also when `is.unsorted(x)`, fixing PR#18294.
- The `data.frame` method of `rbind()` now warns when binding not-wholly-recycling vectors, by analogy to the default method (for matrices).
- `setAs()` finds the correct class for name `to` when multiple packages define a class with that name. Thanks to Gabor Csardi for the report.
- Fix for detaching a package when two classes of the same name are present in method signatures for the same generic. Thanks to Gabor Csardi for the report.
- `match.arg("", c("", "a", "B"))` gives a better error message, in part from PR#17959, thanks to Elin Waring.
- `R CMD Sweave --clean` no longer removes pre-existing files or subdirectories (PR#18242).
- The `quartz()` device no longer splits polylines into subpaths. That has caused narrowly-spaced lines with many points to always look solid even when dashed line type was used due to dash phase restarts.
- Deparsing constructs such as `quote(1 + `!`(2) + 3)` works again as before R 3.5.0, thanks to the report and patch in PR#18284 by Suharto Anggono.
- `as.list(f)` for a `factor f` now keeps `names(f)`, fixing PR#18309.
- `qbeta(.001, .9, .009)` and analogous `qf()` calls now return a correct value instead of `NaN` or wrongly `1`, all with a warning; thanks to the report by Ludger Goeminne in PR#18302.
- `plot.lm()` failed to produce the plot of residuals vs. factor levels (i.e., `which=5` when leverages are constant) for models with character predictors (PR#17840).
- `interaction.plot(..., xtick = TRUE)` misplaced the x-axis line (PR#18305).
- Not strictly fixing a bug, `format()`ing and `print()`ing of non-finite `Date` and `POSIXt` values `NaN` and ±`Inf` no longer show as `NA` but the respective string, e.g., `Inf`, for consistency with numeric vector's behaviour, fulfilling the wish of PR#18308.
- `R CMD check` no longer runs test scripts generated from corresponding '`.Rin`' files twice and now signals an ERROR if processing an '`.Rin`' script fails.
- `tools::Rd2txt()` used for plain-text help pages now renders `\href`s (if `tools::Rd2txt_options(showURLs = TRUE)`) and `\url`s with percent-encoding and standards-compliant delimiting style (angle brackets and no '`URL: `' prefix). `\email` is now rendered with a '`mailto:`' prefix.



CHANGES IN R 4.1.3

NEW FEATURES:

- The default version of Bioconductor has been changed to 3.14. (This is used by `setRepositories` and the menus in GUIs.)

UTILITIES:

- `R CMD check --as-cran` has a workaround for a bug in versions of `file` up to at least 5.41 which mis-identify DBF files last changed in 2022 as executables.

C-LEVEL FACILITIES:

- The legacy S-compatibility macros `SINGLE_*` in '`R_ext/Constants.h`' (included by '`R.h`') are deprecated and will be removed in R 4.2.0.

BUG FIXES:

- Initialization of self-starting `nls()` models with initialization functions following the pre-R-4.1.0 API (without the `...` argument) works again for now, with a deprecation warning.
- Fixed quoting of `~autodetect~` in Java setting defaults to avoid inadvertent user lookup due to leading `~`, reported in PR#18231 by Harold Gutch.
- `substr(., start, stop) <- v` now treats *negative* `stop` values correctly. Reported with a patch in PR#18228 by Brodie Gaslam.
- Subscripting an array `x` without dimnames by a `length(dim(x))`-column character matrix gave "random" non-sense, now an error; reported in PR#18244 by Mikael Jagan.
- `...names()` now matches `names(list(...))` closely, fixing PR#18247.
- `all.equal(*, scale = s)` now works as intended when `length(s) > 1`, partly thanks to Michael Chirico's PR#18272.
- `print(x)` for long vectors `x` now also works for named atomic vectors or lists and prints the correct number when reaching the `getOption("max.print")` limit; partly thanks to a report and proposal by Hugh Parsonage to the R-devel list.
- `all.equal(<selfStart>, *)` no longer signals a deprecation warning.
- `reformulate(*, response=r)` gives a helpful error message now when `length(r) > 1`, thanks to Bill Dunlap's PR#18281.
- Modifying `globalCallingHandlers` inside `withCallingHandlers()` now works or fails correctly, thanks to Henrik Bengtsson's PR#18257.
- `hist(<Date>, breaks = "days")` and `hist(<POSIXt>, breaks = "secs")` no longer fail for inputs of length 1.
- `qbeta(.001, .9, .009)` and similar cases now converge correctly thanks to Ben Bolker's report in PR#17746.
- `window(x, start, end)` no longer wrongly signals "'start' cannot be after 'end'", fixing PR#17527 and PR#18291.
- `data()` now checks that its (rarely used) `list` argument is a character vector – a couple of packages passed other types and gave incorrect results.
- `which()` now checks its `arr.ind` argument is TRUE rather coercing to logical and taking the first element – which gave incorrect results in package code.
- `model.weights()` and `model.offset()` more carefully extract their model components, thanks to Ben Bolker and Tim Taylor's R-devel post.
- `list.files(recursive = TRUE)` now shows all broken symlinks (previously, some of them may have been omitted, PR#18296).



CHANGES IN R 4.1.2

C-LEVEL FACILITIES:

- The workaround in headers '`R.h`' and '`Rmath.h`' (`using namespace std;`) for the Oracle Developer Studio compiler is no longer needed now C++11 is required so has been removed. A couple more usages of `log()` (which should have been `std::log()`) with an `int` argument are reported on Solaris.
- The undocumented limit of 4095 bytes on messages from the S-compatibility macros `PROBLEM` and `MESSAGE` is now documented and longer messages will be silently truncated rather than potentially causing segfaults.
- If the `R_NO_SEGV_HANDLER` environment variable is non-empty, the signal handler for SEGV/ILL/BUS signals (which offers recovery user interface) is not set. This allows more reliable debugging of crashes that involve the console.

DEPRECATED AND DEFUNCT:

- The legacy S-compatibility macros `PROBLEM`, `MESSAGE`, `ERROR`, `WARN`, `WARNING`, `RECOVER`, ... are deprecated and will be hidden in R 4.2.0. R's native interface of `Rf_error` and `Rf_warning` has long been preferred.

BUG FIXES:

- `.mapply(F, dots, .)` no longer segfaults when `dots` is not a `list` and uses `match.fun(F)` as always documented; reported by Andrew Simmons in PR#18164.
- `hist(<Date>, ...)` and `hist(<POSIXt>, ...)` no longer pass arguments for `rect()` (such as `col` and `density`) to `axis()`. (Thanks to Sebastian Meyer's PR#18171.)
- `\Sexpr{ch}` now preserves `Encoding(ch)`. (Thanks to report and patch by Jeroen Ooms in PR#18152.)
- Setting the RNG to `"Marsaglia-Multicarry"` e.g., by `RNGkind()`, now warns in more places, thanks to André Gillibert's report and patch in PR#18168.
- `gray(numeric(), alpha=1/2)` no longer segfaults, fixing PR#18183, reported by Till Krenz.
- Fixed `dnbinom(x, size=<very_small>, .., log=TRUE)` regression, reported by Martin Morgan.
- `as.Date.POSIXlt(x)` now keeps `names(x)`, thanks to Davis Vaughan's report and patch in PR#18188.
- `model.response()` now strips an `"AsIs"` class typically, thanks to Duncan Murdoch's report and other discussants in PR#18190.
- `try()` is considerably faster in case of an error and long call, as e.g., from some `do.call()`. Thanks to Alexander Kaever's suggestion posted to R-devel.
- `qqline(y = <object>)` such as `y=I(.)`, now works, see also PR#18190.
- Non-integer `mgp par()` settings are now handled correctly in `axis()` and `mtext()`, thanks to Mikael Jagan and Duncan Murdoch's report and suggestion in PR#18194.
- `formatC(x)` returns length zero `character()` now, rather than `""` when `x` is of length zero, as documented, thanks to Davis Vaughan's post to R-devel.
- `removeSource(fn)` now retains (other) `attributes(fn)`.



CHANGES IN R 4.1.1

NEW FEATURES:

- `require(`*pkg*`, quietly = TRUE)` is quieter and in particular does not warn if the package is not found.

DEPRECATED AND DEFUNCT:

- Use of '`ftp://`' URIs should be regarded as deprecated, with on-going support confined to `method = "libcurl"` and not routinely tested. (Nowadays no major browser supports them.)
- The non-default `method = "internal"` is deprecated for '`http://`' and '`ftp://`' URIs for both `download.file` and `url`.
- On Windows, `method = "wininet"` is deprecated for '`http://`', '`https://`' and '`ftp://`' URIs for both `download.file` and `url`. (A warning is only given for '`ftp://`'.)
  For '`ftp://`' URIs the default method is now `"libcurl"` if available (which it is on CRAN builds).
  `method = "wininet"` remains the default for '`http://`' and '`https://`' URIs but if `libcurl` is available, using `method = "libcurl"` is preferred.

INSTALLATION:

- `make check` now works also without a LaTeX installation. (Thanks to Sebastian Meyer's PR#18103.)

BUG FIXES:

- `make check-devel` works again in an R build configured with '`--without-recommended-packages`'.
- `qnbinom(p, size, mu)` for large `size/mu` is correct now in a range of cases (PR#18095); similarly for the `(size, prob)` parametrization of the negative binomial. Also `qpois()` and `qbinom()` are better and or faster for extreme cases. The underlying C code has been modularized and is common to all four cases of discrete distributions.
- `gap.axis` is now part of the `axis()` arguments which are passed from `bxp()`, and hence `boxplot()`. (Thanks to Martin Smith's report and suggestions in PR#18109.)
- `.First` and `.Last` can again be set from the site profile.
- `seq.int(from, to, *)` and `seq.default(..)` now work better in large range cases where `from-to` is infinite where the two boundaries are finite.
- `all.equal(x,y)` now returns `TRUE` correctly also when several entries of `abs(x)` and `abs(y)` are close to `.Machine$double.xmax`, the largest finite `numeric`.
- `model.frame()` now clears the object bit when removing the `class` attribute of a value via `na.action` (PR#18100).
- `charClass()` now works with multi-character strings on Windows (PR#18104, fixed by Bill Dunlap).
- `encodeString()` on Solaris now works again in Latin-1 encoding on characters represented differently in UTF-8. Support for surrogate pairs on Solaris has been improved.
- `file.show()` on Windows now works with non-ASCII path names representable in the current native encoding (PR#18132).
- Embedded R on Windows can now find R home directory via the registry even when installed only for the current user (PR#18135).
- `pretty(x)` with finite `x` now returns finite values also in the case where the extreme `x` values are close in size to the maximal representable number `.Machine$double.xmax`. Also, it's been tweaked for very small ranges and when a boundary is close (or equal) to zero; e.g., `pretty(c(0,1e-317))` no longer has negative numbers, currently still warning about a very small range, and `pretty(2^-(1024 - 2^-1/(c(24,10))))` is more accurate.

- The error message for not finding vignette files when weaving has correct file sizes now. (Thanks to Sebastian Meyer's PR#18154.)
- `dnbinom(20, <large>, 1)` now correctly gives 0, and similar cases are more accurate with underflow precaution. (Reported by Francisco Vera Alcivar in PR#18072.)



CHANGES IN R 4.1.0

FUTURE DIRECTIONS:

- It is planned that the 4.1.x series will be the last to support 32-bit Windows, with production of binary packages for that series continuing until early 2023.

SIGNIFICANT USER-VISIBLE CHANGES:

- Data set `esoph` in package **datasets** now provides the correct numbers of controls; previously it had the numbers of cases added to these. (Reported by Alexander Fowler in PR#17964.)

NEW FEATURES:

- '`www.omegahat.net`' is no longer one of the repositories known by default to `setRepositories()`. (Nowadays it only provides source packages and is often unavailable.)
- Function `package_dependencies()` (in package **tools**) can now use different dependency types for direct and recursive dependencies.
- The checking of the size of tarball in `R CMD check --as-cran <pkg>` may be tweaked via the new environment variable `_R_CHECK_CRAN_INCOMING_TARBALL_THRESHOLD_`, as suggested in PR#17777 by Jan Gorecki.
- Using `c()` to combine a factor with other factors now gives a factor, an ordered factor when combining ordered factors with identical levels.
- `apply()` gains a `simplify` argument to allow disabling of simplification of results.
- The `format()` method for class `"ftable"` gets a new option `justify`. (Suggested by Thomas Soeiro.)
- New `...names()` utility. (Proposed by Neal Fultz in PR#17705.)
- `type.convert()` now warns when its `as.is` argument is not specified, as the help file always said it *should*. In that case, the default is changed to `TRUE` in line with its change in `read.table()` (related to `stringsAsFactors`) in R 4.0.0.
- When printing list arrays, classed objects are now shown *via* their `format()` value if this is a short enough character string, or by giving the first elements of their class vector and their length.
- `capabilities()` gets new entry `"Rprof"` which is `TRUE` when R has been configured with the equivalent of `--enable-R-profiling` (as it is by default). (Related to Michael Orlitzky's report PR#17836.)
- `str(xS4)` now also shows extraneous attributes of an S4 object `xS4`.
- Rudimentary support for vi-style tags in `rtags()` and `R CMD rtags` has been added. (Based on a patch from Neal Fultz in PR#17214.)
- `checkRdContents()` is now exported from **tools**; it and also `checkDocFiles()` have a new option `chkInternal` allowing to check Rd files marked with keyword `"internal"` as well. The latter can be activated for `R CMD check` via environment variable `_R_CHECK_RD_INTERNAL_TOO_`.
- New functions `numToBits()` and `numToInts()` extend the `raw` conversion utilities to (double precision) `numeric`.

- Functions `URLencode()` and `URLdecode()` in package **utils** now work on vectors of URIs. (Based on patch from Bob Rudis submitted with PR#17873.)
- `path.expand()` can expand '`~user`' on most Unix-alikes even when `readline` is not in use. It tries harder to expand '`~`', for example should environment variable `HOME` be unset.
- For HTML help (both dynamic and static), Rd file links to help pages in external packages are now treated as references to topics rather than file names, and fall back to a file link only if the topic is not found in the target package. The earlier rule which prioritized file names over topics can be restored by setting the environment variable `_R_HELP_LINKS_TO_TOPICS_` to a false value.
- `c()` now removes `NULL` arguments before dispatching to methods, thus simplifying the implementation of `c()` methods, *but* for back compatibility keeps `NULL` when it is the first argument. (From a report and patch proposal by Lionel Henry in PR#17900.)
- `Vectorize()`'s result function's environment no longer keeps unneeded objects.
- Function `...elt()` now propagates visibility consistently with `..n`. (Thanks to Lionel Henry's PR#17905.)
- `capture.output()` no longer uses non-standard evaluation to evaluate its arguments. This makes evaluation of functions like `parent.frame()` more consistent. (Thanks to Lionel Henry's PR#17907.)
- `packBits(bits, type="double")` now works as inverse of `numToBits()`. (Thanks to Bill Dunlap's proposal in PR#17914.)
- `curlGetHeaders()` has two new arguments, `timeout` to specify the timeout for that call (overriding `getOption("timeout")`) and `TLS` to specify the minimum TLS protocol version to be used for `https://` URIs (*inter alia* providing a means to check for sites using deprecated TLS versions 1.0 and 1.1).
- For `nls()`, an optional constant `scaleOffset` may be added to the denominator of the relative offset convergence test for cases where the fit of a model is expected to be exact, thanks to a proposal by John Nash. `nls(*, trace=TRUE)` now also shows the convergence criterion.
- Numeric differentiation *via* `numericDeriv()` gets new optional arguments `eps` and `central`, the latter for taking central divided differences. The latter can be activated for `nls()` via `nls.control(nDcentral = TRUE)`.
- `nls()` now passes the `trace` and `control` arguments to `getInitial()`, notably for all self-starting models, so these can also be fit in zero-noise situations via a `scaleOffset`. For this reason, the `initial` function of a `selfStart` model must now have `...` in its argument list.
- `bquote(splice = TRUE)` can now splice expression vectors with attributes: this makes it possible to splice the result of `parse(keep.source = TRUE)`. (Report and patch provided by Lionel Henry in PR#17869.)
- `textConnection()` gets an optional `name` argument.
- `get()`, `exists()`, and `get0()` now signal an error if the first argument has length greater than 1. Previously additional elements were silently ignored. (Suggested by Antoine Fabri on R-devel.)
- R now provides a shorthand notation for creating functions, e.g. `\(x) x + 1` is parsed as `function(x) x + 1`.
- R now provides a simple native forward pipe syntax `|>`. The simple form of the forward pipe inserts the left-hand side as the first argument in the right-hand side call. The pipe implementation as a syntax transformation was motivated by suggestions from Jim Hester and Lionel Henry.

- `all.equal(f, g)` for `functions` now by default also compares their `environment(.)`s, notably via new `all.equal` method for class `function`. Comparison of `nls()` fits, e.g., may now need `all.equal(m1, m2,check.environment = FALSE)`.
- `.libPaths()` gets a new option `include.site`, allowing to *not* include the site library. (Thanks to Dario Strbenac's suggestion and Gabe Becker's PR#18016.)
- Lithuanian translations are now available. (Thanks to Rimantas Žakauskas.)
- `names()` now works for `DOTSXP` objects. On the other hand, in 'R-lang', the R language manual, we now warn against relying on the structure or even existence of such dot-dot-dot objects.
- `all.equal()` no longer gives an error on `DOTSXP` objects.
- `capabilities("cairo")` now applies only to the file-based devices as it is now possible (if very unusual) to build R with Cairo support for those but not for `X11()`.
- There is optional support for tracing the progress of `loadNamespace()` — see its help.
- (Not Windows.) `l10n_info()` reports an additional element, the name of the encoding as reported by the OS (which may differ from the encoding part (if any) of the result from `Sys.getlocale("LC_CTYPE")`).
- New function `gregexec()` which generalizes `regexec()` to find *all* disjoint matches and all substrings corresponding to parenthesized subexpressions of the given regular expression. (Contributed by Brodie Gaslam.)
- New function `charClass()` in package **utils** to query the wide-character classification functions in use (such as `iswprint`).
- The names of `quantile()`'s result no longer depend on the global `getOption("digits")`, but `quantile()` gets a new optional argument `digits = 7` instead.
- `grep()`, `sub()`, `regexp` and variants work considerably faster for long factors with few levels. (Thanks to Michael Chirico's PR#18063.)
- Provide grouping of `x11()` graphics windows within a window manager such as `Gnome` or `Unity`; thanks to a patch by Ivan Krylov posted to R-devel.
- The `split()` method for class `data.frame` now allows the `f` argument to be specified as a formula.
- `sprintf` now warns on arguments unused by the format string.
- New palettes `"Rocket"` and `"Mako"` for `hcl.colors()` (approximating palettes of the same name from the **viridisLite** package).
  Contributed by Achim Zeileis.
- The base environment and its namespace are now locked (so one can no longer add bindings to these or remove from these).
- `Rterm` handling of multi-byte characters has been improved, allowing use of such characters when supported by the current locale.
- `Rterm` now accepts `ALT+ +xxxxxxxx` sequences to enter Unicode characters as hex digits.
- Environment variable `LC_ALL` on Windows now takes precedence over `LC_CTYPE` and variables for other supported categories, matching the POSIX behaviour.
- `duplicated()` and `anyDuplicated()` are now optimized for integer and real vectors that are known to be sorted via the ALTREP framework. Contributed by Gabriel Becker via PR#17993.

GRAPHICS:

- The graphics engine version, `R_GE_version`, has been bumped to `14` and so packages that provide graphics devices should be reinstalled.

- Graphics devices should now specify `deviceVersion` to indicate what version of the graphics engine they support.
- Graphics devices can now specify `deviceClip`. If `TRUE`, the graphics engine will never perform any clipping of output itself.
  The clipping that the graphics engine does perform (for both `canClip = TRUE` and `canClip = FALSE`) has been improved to avoid producing unnecessary artifacts in clipped output.
- The **grid** package now allows `gpar(fill)` to be a `linearGradient()`, a `radialGradient()`, or a `pattern()`. The `viewport(clip)` can now also be a grob, which defines a clipping path, and there is a new `viewport(mask)` that can also be a grob, which defines a mask.
  These new features are only supported so far on the Cairo-based graphics devices and on the `pdf()` device.
- (Not Windows.) A warning is given when a Cairo-based type is specified for a `png()`, `jpeg()`, `tiff()` or `bmp()` device but Cairo is unsupported (so `type = "Xlib"` is tried instead).
- `grSoftVersion()` now reports the versions of FreeType and FontConfig if they are used directly (not *via* Pango), as is most commonly done on macOS.

C-LEVEL FACILITIES:

- The *standalone* '`libRmath`' math library and R's C API now provide `log1pexp()` again as documented, and gain `log1mexp()`.

INSTALLATION on a UNIX-ALIKE:

- `configure` checks for a program `pkgconf` if program `pkg-config` is not found. These are now only looked for on the path (like almost all other programs) so if needed specify a full path to the command in `PKG_CONFIG`, for example in file '`config.site`'.
- C99 function `iswblank` is required – it was last seen missing ca 2003 so the workaround has been removed.
- There are new `configure` options '`--with-internal-iswxxxxx`', '`--with-internal-towlower`' and '`--with-internal-wcwidth`' which allows the system functions for wide-character classification, case-switching and width (`wcwidth` and `wcswidth`) to be replaced by internal ones. The first has long been used on macOS, AIX (and Windows) but this enables it to be unselected there and selected for other platforms (it is the new default on Solaris). The second is new in this version of R and is selected by default on macOS and Solaris. The third has long been the default and remains so as it contains customizations for East Asian languages.
  System versions of these functions are often minimally implemented (sometimes only for ASCII characters) and may not cover the full range of Unicode points: for example Solaris (and Windows) only cover the Basic Multilingual Plane.
- Cairo installations without X11 are more likely to be detected by `configure`, when the file-based Cairo graphics devices will be available but not `X11(type = "cairo")`.
- There is a new `configure` option '`--with-static-cairo`' which is the default on macOS. This should be used when only static cairo (and where relevant, Pango) libraries are available.
- Cairo-based graphics devices on platforms without Pango but with FreeType/FontConfig will make use of the latter for font selection.

LINK-TIME OPTIMIZATION on a UNIX-ALIKE:

- Configuring with flag '`--enable-lto=R`' now also uses LTO when installing the recommended packages.

- `R CMD INSTALL` and `R CMD SHLIB` have a new flag '`--use-LTO`' to use LTO when compiling code, for use with R configured with '`--enable-lto=R`'. For R configured with '`--enable-lto`', they have the new flag '`--no-use-LTO`'.
  Packages can opt in or out of LTO compilation *via* a '`UseLTO`' field in the '`DESCRIPTION`' file. (As usual this can be overridden by the command-line flags.)

BUILDING R on Windows:

- for GCC $\geq$ 8, `FC_LEN_T` is defined in '`config.h`' and hence character lengths are passed from C to Fortran in *inter alia* BLAS and LAPACK calls.
- There is a new text file '`src/gnuwin32/README.compilation`', which outlines how C/Fortran code compilation is organized and documents new features:
  - R can be built with Link-Time Optimization with a suitable compiler – doing so with GCC 9.2 showed several inconsistencies which have been corrected.
  - There is support for cross-compiling the C and Fortran code in R and standard packages on suitable (Linux) platforms. This is mainly intended to allow developers to test later versions of compilers – for example using GCC 9.2 or 10.x has detected issues that GCC 8.3 in Rtools40 does not.
  - There is experimental support for cross-building R packages with C, C++ and/or Fortran code.
- The R installer can now be optionally built to support a single architecture (only 64-bit or only 32-bit).

PACKAGE INSTALLATION:

- The default C++ standard has been changed to C++14 where available (which it is on all currently checked platforms): if not (as before) C++11 is used if available otherwise C++ is not supported.
  Packages which specify C++11 will still be installed using C++11.
  C++14 compilers may give deprecation warnings, most often for `std::random_shuffle` (deprecated in C++14 and removed in C++17). Either specify C++11 (see 'Writing R Extensions') or modernize the code and if needed specify C++14. The latter has been supported since R 3.4.0 so the package's '`DESCRIPTION`' would need to include something like

  ```
  Depends: R (>= 3.4)
  ```

PACKAGE INSTALLATION on Windows:

- `R CMD INSTALL` and `R CMD SHLIB` make use of their flag '`--use-LTO`' when the '`LTO_OPT`' make macro is set in file '`etc/${R_ARCH}/Makeconf`' or in a personal/site '`Makevars`' file. (For details see 'Writing R Extensions' §4.5.)
  This provides a valuable check on code consistency. It does work with GCC 8.3 as in Rtools40, but that does not detect everything the CRAN checks with current GCC do.

PACKAGE INSTALLATION on macOS:

- The default personal library directory on builds with '`--enable-aqua`' (including CRAN builds) now differs by CPU type, one of

  ```
  ~/Library/R/x86_64/x.y/library
  ~/Library/R/arm64/x.y/library
  ```

  This uses the CPU type R (and hence the packages) were built for, so when a '`x86_64`' build of R is run under Rosetta emulation on an '`arm64`' Mac, the first is used.

UTILITIES:

- `R CMD check` can now scan package functions for bogus `return` statements, which were possibly intended as `return()` calls (wish of PR#17180, patch by Sebastian Meyer). This check can be activated via the new environment variable `_R_CHECK_BOGUS_RETURN_`, true for `--as-cran`.
- `R CMD build` omits tarballs and binaries of previous builds from the top-level package directory. (PR#17828, patch by Sebastian Meyer.)
- `R CMD check` now runs sanity checks on the use of '`LazyData`', for example that a '`data`' directory is present and that '`LazyDataCompression`' is not specified without '`LazyData`' and has a documented value. For packages with large LazyData databases without specifying '`LazyDataCompression`', there is a reference to the code given in 'Writing R Extensions' §1.1.6 to test the choice of compression (as in all the CRAN packages tested a non-default method was preferred).
- `R CMD build` removes '`LazyData`' and '`LazyDataCompression`' fields from the '`DESCRIPTION`' file of packages without a '`data`' directory.

ENCODING-RELATED CHANGES:

- The parser now treats '`\Unnnnnnnn`' escapes larger than the upper limit for Unicode points ('`\U10FFFF`') as an error as they cannot be represented by valid UTF-8.
  Where such escapes are used for outputting non-printable (including unassigned) characters, 6 hex digits are used (rather than 8 with leading zeros). For clarity, braces are used, for example '`\U{0effff}`'.
- The parser now looks for non-ASCII spaces on Solaris (as previously on most other OSes).
- There are warnings (including from the parser) on the use of unpaired surrogate Unicode points such as '`\uD834`'. (These cannot be converted to valid UTF-8.)
- Functions `nchar()`, `tolower()`, `toupper()` and `chartr()` and those using regular expressions have more support for inputs with a marked Latin-1 encoding.
- The character-classification functions used (by default) to replace the system `iswxxxxx` functions on Windows, macOS and AIX have been updated to Unicode 13.0.0.
  The character-width tables have been updated to include new assignments in Unicode 13.0.0. This included treating all control characters as having zero width.
- The code for evaluating default (extended) regular expressions now uses the same character-classification functions as the rest of R (previously they differed on Windows, macOS and AIX).
- There is a build-time option to replace the system's wide-character `wctrans` C function by tables shipped with R: use `configure` option '`--with-internal-towlower`' or (on Windows) '`-DUSE_RI18N_CASE`' in '`CFLAGS`' when building R. This may be needed to allow `tolower()` and `toupper()` to work with Unicode characters beyond the Basic Multilingual Plane where not supported by system functions (e.g. on Solaris where it is the new default).
- R is more careful when truncating UTF-8 and other multi-byte strings that are too long to be printed, passed to the system or libraries or placed into an internal buffer. Truncation will no longer produce incomplete multibyte characters.

DEPRECATED AND DEFUNCT:

- Function `plclust()` from the package **stats** and `package.dependencies()`, `pkgDepends()`, `getDepList()`, `installFoundDepends()`, and `vignetteDepends()` from package **tools** are defunct.
- Defunct functions `checkNEWS()` and `readNEWS()` from package **tools** and `CRAN.packages()` from **utils** have been removed.

- `R CMD config CXXCPP` is defunct (it was deprecated in R 3.6.2).
- `parallel::detectCores()` drops support for Irix (retired in 2013).
- The `LINPACK` argument to `chol.default()`, `chol2inv()`, `solve.default()` and `svd()` has been defunct since R 3.1.0. It was silently ignored up to R 4.0.3 but now gives an error.
- Subsetting/indexing, such as `ddd[*]` or `ddd$x` on a `DOTSXP` (dot-dot-dot) object `ddd` has been disabled; it worked by accident only and was undocumented.

BUG FIXES:

- Many more C-level allocations (mainly by `malloc` and `strdup`) are checked for success with suitable alternative actions.
- Bug fix for `replayPlot()`; this was turning off graphics engine display list recording if a recorded plot was replayed in the same session. The impact of the bug became visible if resize the device after replay OR if attempted another `savePlot()` after replay (empty display list means empty screen on resize or empty saved plot).
- `R CMD check` etc now warn when a package exports non-existing S4 classes or methods, also in case of no "methods" presence. (Reported by Alex Bertram; reproducible example and patch by Sebastian Meyer in PR#16662.)
- `boxplot()` now also accepts `call`s for labels such as `ylab`, the same as `plot()`. (Reported by Marius Hofert.)
- The help page for `xtabs()` now correctly states that `addNA` is setting `na.action = na.pass` among others. (Reported as PR#17770 by Thomas Soeiro.)
- The `R CMD check <pkg>` gives a longer and more comprehensible message when '`DESCRIPTION`' misses dependencies, e.g., in `Imports:`. (Thanks to the contributors of PR#17179.)
- `update.default()` now calls the generic `update()` on the formula to work correctly for models with extended formulas. (As reported and suggested by Neal Fultz in PR#17865.)
- The horizontal position of leaves in a dendrogram is now correct also with `center = FALSE`. (PR#14938, patch from Sebastian Meyer.)
- `all.equal.POSIXt()` no longer warns about and subsequently ignores inconsistent `"tzone"` attributes, but describes the difference in its return value (PR#17277). This check can be disabled *via* the new argument `check.tzone = FALSE` as suggested by Sebastian Meyer.
- `as.POSIXct()` now populates the `"tzone"` attribute from its `tz` argument when `x` is a logical vector consisting entirely of `NA` values.
- `x[[2^31]] <- v` now works. (Thanks to the report and patch by Suharto Anggono in PR#17330.)
- In log-scale graphics, `axis()` ticks and label positions are now computed more carefully and symmetrically in their range, typically providing *more* ticks, fulfilling wishes in PR#17936. The change really corresponds to an improved `axisTicks()` (package **grDevices**), potentially influencing **grid** and **lattice**, for example.
- `qnorm(<very large negative>, log.p=TRUE)` is now correct to at least five digits where it was catastrophically wrong, previously.
- `sum(df)` and similar `"Summary"`- and `"Math"`-group member functions now work for data frames `df` with `logical` columns, notably also of zero rows. (Reported to R-devel by Martin "b706".)
- `unsplit()` had trouble with tibbles due to unsound use of `rep(NA, len)`-indexing, which should use `NA_integer_` (Reported to R-devel by Mario Annau.)
- `pnorm(x, log.p = TRUE)` underflows to `-Inf` slightly later.

- `show(<hidden S4 generic>)` prints better and without quotes for non-hidden S4 generics.
- `read.table()` and relatives treated an "NA" column name as missing when `check.names = FALSE` PR#18007.
- Parsing strings containing UTF-16 surrogate pairs such as `"\uD834\uDD1E"` works better on some (uncommon) platforms. `sprintf("%X", utf8ToInt("\uD834\uDD1E"))` should now give `"1D11E"` on all platforms.
- `identical(x,y)` is no longer true for differing `DOTSXP` objects, fixing PR#18032.
- `str()` now works correctly for `DOTSXP` and related exotics, even when these are doomed.
  Additionally, it no longer fails for `lists` with a `class` and "irregular" method definitions such that e.g. `lapply(*)` will necessarily fail, as currently for different **igraph** objects.
- Message translation domains, e.g., for errors and warnings, are now correctly determined also when e.g., a **base** function is called from "top-level" function (i.e., defined in `globalenv()`), thanks to a patch from Joris Goosen fixing PR#17998.
- Too long lines in environment files (e.g., '`Renviron`') no longer crash R. This limit has been increased to 100,000 bytes. (PR#18001.)
- There is a further workaround for FreeType giving incorrect italic font faces with cairo-based graphics devices on macOS.
- `add_datalist(*, force = TRUE)` (from package **tools**) now actually updates an existing '`data/datalist`' file for new content. (Thanks to a report and patch by Sebastian Meyer in PR#18048.)
- `cut.Date()` and `cut.POSIXt()` could produce an empty last interval for `breaks = "months"` or `breaks = "years"`. (Reported as PR#18053 by Christopher Carbone.)
- Detection of the encoding of 'regular' macOS locales such as '`en_US`' (which is UTF-8) had been broken by a macOS change: fortunately these are now rarely used with '`en_US.UTF-8`' being preferred.
- `sub()` and `gsub(pattern, repl, x, *)` now keep attributes of `x` such as `names()` also when `pattern` is `NA` (PR#18079).
- Time differences (`"difftime"` objects) get a replacement and a `rep()` method to keep `"units"` consistent. (Thanks to a report and patch by Nicolas Bennett in PR#18066.)
- The `\RdOpts` macro, setting defaults for `\Sexpr` options in an Rd file, had been ineffective since R 2.12.0: it now works again. (Thanks to a report and patch by Sebastian Meyer in PR#18073.)
- `mclapply` and `pvec` no longer accidentally terminate parallel processes started before by `mcparallel` or related calls in package **parallel** (PR#18078).
- `grep` and other functions for evaluating (extended) regular expressions handle in Unicode also strings not explicitly flagged UTF-8, but flagged native when running in UTF-8 locale.
- Fixed a crash in `fifo` implementation on Windows (PR#18031).
- Binary mode in `fifo` on Windows is now properly detected from argument `open` (PR#15600, PR#18031).

CHANGES IN R 4.0.5

BUG FIXES:

- The change to the internal table in R 4.0.4 for `iswprint` has been reverted: it contained some errors in printability of 'East Asian' characters.

- For packages using 'LazyData', R CMD build ignored the '--resave-data' option and the 'BuildResaveData' field of the 'DESCRIPTION' file (in R versions 4.0.0 to 4.0.4).



CHANGES IN R 4.0.4

NEW FEATURES:

- File 'share/texmf/tex/latex/jss.cls' has been updated to work with LaTeX versions since Oct 2020.
- Unicode character width tables (as used by nchar(, type = "w")) have been updated to Unicode 12.1 by Brodie Gaslam (PR#17781), including many emoji.
- The internal table for iswprint (used on Windows, macOS and AIX) has been updated to include many recent Unicode characters.

INSTALLATION on a UNIX-ALIKE:

- If an external BLAS is specified by '--with-blas=foo' or *via* environment variable BLAS_LIBS is not found, this is now a configuration error. The previous behaviour was not clear from the documentation: it was to continue the search as if '--with-blas=yes' was specified.

BUG FIXES:

- all.equal(x,y) now "sees" the two different NAs in factors, thanks to Bill Dunlap and others in PR#17897.
- (~ NULL)[1] and similar formula subsetting now works, thanks to a report and patch by Henrik Bengtsson in PR#17935. Additionally, subsetting leaving an empty formula now works too, thanks to suggestions by Suharto Anggono.
- .traceback(n) keeps source references again, as before R 4.0.0, fixing a regression; introduced by the PR#17580, reported including two patch proposals by Brodie Gaslam.
- unlist(plst, recursive=FALSE) no longer drops content for pairlists with list components, thanks to the report and patch by Suharto Anggono in PR#17950.
- iconvlist() now also works on MUSL based (Linux) systems, from a report and patch suggestion by Wesley Chan in PR#17970.
- round() and signif() no longer tolerate wrong argument names, notably in 1-argument calls; reported by Shane Mueller on R-devel (mailing list); later reported as PR#17976.
- .Machine has longdouble.* elements only if capabilities("long.double") is true, as documented. (Previously they were included if the platform had long double identical to double, as ARM does.)
- p.adjust(numeric(), n=0) now works, fixing PR#18002.
- identical(x,y) no longer prints "Unknown Type .." for typeof(x) == "..." objects.
- Fix (auto-)print()ing of named complex vectors, see PR#17868 and PR#18019.
- all.equal(<language>, <...>) now works, fixing PR#18029.
- as.data.frame.list(L, row.names=NULL) now behaves in line with data.frame(), disregarding names of components of L, fixing PR#18034, reported by Kevin Tappe.
- checkRdaFiles(ff)$version is now correct also when ff contains files of different versions, thanks to a report and patch from Sebastian Meyer in PR#18041.
- macOS: Quartz device live drawing could fail (no plot is shown) if the system changes the drawing context after view update (often the case since macOS Big Sur). System log may show "CGContextDelegateCreateForContext: invalid context" error.



CHANGES IN R 4.0.3

NEW FEATURES:

- On platforms using `configure` option '`--with-internal-tzcode`', additional values `"internal"` and (on macOS only) `"macOS"` are accepted for the environment variable `TZDIR`. (See `?TZDIR`.)
  On macOS, `"macOS"` is used by default if the system timezone database is a newer version than that in the R installation.
- When `install.packages(type = "source")` fails to find a package in a repository it mentions package versions which are excluded by their R version requirement and links to hints on why a package might not be found.
- The default value for `options("timeout")` can be set from environment variable `R_DEFAULT_INTERNET_TIMEOUT`, still defaulting to 60 (seconds) if that is not set or invalid.
  This may be needed when child R processes are doing downloads, for example during the installation of source packages which download jars or other forms of data.

LINK-TIME OPTIMIZATION on a UNIX-ALIKE:

- There is now support for parallelized Link-Time Optimization (LTO) with GCC and for 'thin' LTO with `clang` *via* setting the '`LTO`' macro.
- There is support for setting a different LTO flag for the Fortran compiler, including to empty when mixing `clang` and `gfortran` (as on macOS). See file '`config.site`'.
- There is a new '`LTO_LD`' macro to set linker options for LTO compilation, for example to select an alternative linker or to parallelize thin LTO.

DEPRECATED AND DEFUNCT:

- The `LINPACK` argument to `chol.default()`, `chol2inv()`, `solve.default()` and `svd()` has been defunct since R 3.1.0. Using it now gives a warning which will become an error in R 4.1.0.

BUG FIXES:

- The code mitigating stack overflow with PCRE regexps on very long strings is enabled for PCRE2 < 10.30 also when JIT is enabled, since stack overflows have been seen in that case.
- Fix to correctly show the group labels in `dotchart()` (which where lost in the `ylab` improvement for R 4.0.0).
- `addmargins(*, ..)` now also works when `fn()` is a local function, thanks to bug report and patch PR#17124 from Alex Bertram.
- `rank(x)` and hence `sort(x)` now work when `x` is an object (as per `is.object(x)`) of type `"raw"` *and* provides a valid `` `[` `` method, e.g., for `gmp::as.bigz(.)` numbers.
- `chisq.test(*, simulate.p.value=TRUE)` and `r2dtable()` now work correctly for large table entries (in the millions). Reported by Sebastian Meyer and investigated by more helpers in PR#16184.
- Low-level socket read/write operations have been fixed to correctly signal communication errors. Previously, such errors could lead to a segfault due to invalid memory access. Reported and debugged by Dmitriy Selivanov in PR#17850.
- `quantile(x, pr)` works more consistently for `pr` values slightly outside [0,1], thanks to Suharto Anggono's PR#17891.
  Further, `quantile(x, prN, names=FALSE)` now works even when `prN` contains `NA`s, thanks to Anggono's PR#17892. Ditto for ordered factors or `Date` objects when `type = 1` or `3`, thanks to PR#17899.

- Libcurl-based internet access, including `curlGetHeaders()`, was not respecting the `"timeout"` option. If this causes unanticipated timeouts, consider increasing the default by setting `R_DEFAULT_INTERNET_TIMEOUT`.
- `as.Date(<char>)` now also works with an initial `""`, thanks to Michael Chirico's PR#17909.
- `isS3stdGeneric(f)` now detects an S3 generic also when it is `trace()`d, thanks to Gabe Becker's PR#17917.
- `R_allocLD()` has been fixed to return memory aligned for long double type PR#16534.
- `fisher.test()` no longer segfaults when called again after its internal stack has been exceeded PR#17904.
- Accessing a long vector represented by a compact integer sequence no longer segfaults (reported and debugged by Hugh Parsonage).
- `duplicated()` now works also for strings with multiple encodings inside a single vector PR#17809.
- `phyper(11, 15, 0, 12, log.p=TRUE)` no longer gives `NaN`; reported as PR#17271 by Alexey Stukalov.
- Fix incorrect calculation in `logLik.nls()` PR#16100, patch from Sebastian Meyer.
- A very old bug could cause a segfault in `model.matrix()` when terms involved logical variables. Part of PR#17879.
- `model.frame.default()` allowed `data = 1`, leading to involuntary variable capture (rest of PR#17879).
- `tar()` no longer skips non-directory files, thanks to a patch by Sebastian Meyer, fixing the remaining part of PR#16716.



CHANGES IN R 4.0.2

UTILITIES:

- `R CMD check` skips vignette re-building (with a warning) if the '`VignetteBuilder`' package(s) are not available.

BUG FIXES:

- Paths with non-ASCII characters caused problems for package loading on Windows PR#17833.
- Using **tcltk** widgets no longer crashes R on Windows.
- `source(*, echo=TRUE)` no longer fails in some cases with empty lines; reported by Bill Dunlap in PR#17769.
- `on.exit()` now correctly matches named arguments, thanks to PR#17815 (including patch) by Brodie Gaslam.
- `regexpr(*, perl=TRUE)` no longer returns incorrect positions into text containing characters outside of the Unicode Basic Multilingual Plane on Windows.



CHANGES IN R 4.0.1

NEW FEATURES:

- `paste()` and `paste0()` gain a new optional argument `recycle0`. When set to true, zero-length arguments are recycled leading to `character(0)` after the `sep`-concatenation, i.e., to the empty string `""` if `collapse` is a string and to the zero-length value `character(0)` when `collapse = NULL`.
  A package whose code uses this should depend on '`R (>= 4.0.1)`'.
- The `summary(<warnings>)` method now maps the counts correctly to the warning messages.

BUG FIXES:

- `aov(frml, ...)` now also works where the `formula` deparses to more than 500 characters, thanks to a report and patch proposal by Jan Hauffa.
- Fix a dozen places (code, examples) as `Sys.setlocale()` returns the new rather than the previous setting.
- Fix for adding two complex **grid** units via `sum()`. Thanks to Gu Zuguang for the report and Thomas Lin Pedersen for the patch.
- Fix `parallel::mclapply(..., mc.preschedule=FALSE)` to handle raw vector results correctly. PR#17779
- Computing the `base` value, i.e., 2, "everywhere", now uses `FLT_RADIX`, as the original '`machar`' code looped indefinitely on the ppc64 architecture for the `longdouble` case.
- In R 4.0.0, `sort.list(x)` when `is.object(x)` was true, e.g., for `x <- I(letters)`, was accidentally using `method = "radix"`. Consequently, e.g., `merge(<data.frame>)` was much slower than previously; reported in PR#17794.
- `plot(y ~ x, ylab = quote(y[i]))` now works, as e.g., for `xlab`; related to PR#10525.
- `parallel::detect.cores(all.tests = TRUE)` tries a matching OS name before the other tests (which were intended only for unknown OSes).
- Parse data for raw strings is now recorded correctly. Reported by Gabor Csardi.

CHANGES IN R 4.0.0

SIGNIFICANT USER-VISIBLE CHANGES:

- Packages need to be (re-)installed under this version (4.0.0) of R.
- `matrix` objects now also inherit from class `"array"`, so e.g., `class(diag(1))` is `c("matrix","array")`. This invalidates code incorrectly assuming that `class(matrix_obj))` has length one.
  S3 methods for class `"array"` are now dispatched for `matrix` objects.
- There is a new syntax for specifying *raw* character constants similar to the one used in C++: `r"(...)"` with `...` any character sequence not containing the sequence '`)"`'. This makes it easier to write strings that contain backslashes or both single and double quotes. For more details see `?Quotes`.
- R now uses a '`stringsAsFactors = FALSE`' default, and hence by default no longer converts strings to factors in calls to `data.frame()` and `read.table()`.
  A large number of packages relied on the previous behaviour and so have needed/will need updating.
- The `plot()` S3 generic function is now in package **base** rather than package **graphics**, as it is reasonable to have methods that do not use the **graphics** package. The generic is currently re-exported from the **graphics** namespace to allow packages importing it from there to continue working, but this may change in future.

Packages which define S4 generics for `plot()` should be re-installed and package code using such generics from other packages needs to ensure that they are imported rather than rely on their being looked for on the search path (as in a namespace, the base namespace has precedence over the search path).

REFERENCE COUNTING:

- Reference counting is now used instead of the `NAMED` mechanism for determining when objects can be safely mutated in base C code. This reduces the need for copying in some cases and should allow further optimizations in the future. It should help make the internal code easier to maintain.
  This change is expected to have almost no impact on packages using supported coding practices in their C/C++ code.

MIGRATION TO PCRE2:

- This version of R is built against the PCRE2 library for Perl-like regular expressions, if available. (On non-Windows platforms PCRE1 can optionally be used if PCRE2 is not available at build time.) The version of PCRE in use can be obtained *via* `extSoftVersion()`: PCRE1 (formerly known as 'PCRE') has versions <= 8, PCRE2 versions >= 10.
- Making PCRE2 available when building R from source is strongly recommended (preferably version 10.30 or later) as PCRE1 is no longer developed: version 8.44 is 'likely to be the final release'.
- PCRE2 reports errors for some regular expressions that were accepted by PCRE1. A hyphen now has to be escaped in a character class to be interpreted as a literal (unless first or last in the class definition). '`\R`', '`\B`' and '`\X`' are no longer allowed in character classes (PCRE1 treated these as literals).
- Option `PCRE_study` is no longer used with PCRE2, and is reported as `FALSE` when that is in use.

NEW FEATURES:

- `assertError()` and `assertWarning()` (in package **tools**) can now check for *specific* error or warning classes *via* the new optional second argument `classes` (which is not back compatible with previous use of an unnamed second argument).
- `DF2formula()`, the utility for the data frame method of `formula()`, now works without parsing and explicit evaluation, starting from Suharto Anggono's suggestion in PR#17555.
- `approxfun()` and `approx()` gain a new argument `na.rm` defaulting to true. If set to false, missing `y` values now propagate into the interpolated values.
- Long vectors are now supported as the `seq` argument of a `for()` loop.
- `str(x)` gets a new `deparse.lines` option with a default to speed it up when `x` is a large `call` object.
- The internal traceback object produced when an error is signalled (`.Traceback`), now contains the `call`s rather than the `deparse()`*d* calls, deferring the deparsing to the user-level functions `.traceback()` and `traceback()`. This fulfils the wish of PR#17580, reported including two patch proposals by Brodie Gaslam.
- `data.matrix()` now converts character columns to factors and from this to integers.
- `package.skeleton()` now explicitly lists all exports in the '`NAMESPACE`' file.
- New function `.S3method()` to register S3 methods in R scripts.
- `file.path()` has some support for file paths not in the session encoding, e.g. with UTF-8 inputs in a non-UTF-8 locale the output is marked as UTF-8.

- Most functions with file-path inputs will give an explicit error if a file-path input in a marked encoding cannot be translated (to the native encoding or in some cases on Windows to UTF-8), rather than translate to a different file path using escapes. Some (such as `dir.exists()`, `file.exists()`, `file.access()`, `file.info()`, `list.files()`, `normalizePath()` and `path.expand()`) treat this like any other non-existent file, often with a warning.
- There is a new help document accessed by `help("file path encoding")` detailing how file paths with marked encodings are handled.
- New function `list2DF()` for creating data frames from lists of variables.
- `iconv()` has a new option `sub = "Unicode"` to translate UTF-8 input invalid in the '`to`' encoding using '`<U+xxxx>`' escapes.
- There is a new function `infoRDS()` providing information about the serialization format of a serialized object.
- S3 method lookup now by default skips the elements of the search path between the global and base environments.
- Added an argument `add_datalist(*, small.size = 0)` to allow the creation of a '`data/datalist`' file even when the total size of the data sets is small.
- The backquote function `bquote()` has a new argument `splice` to enable splicing a computed list of values into an expression, like `,@` in LISP's backquote.
- The formula interface to `t.test()` and `wilcox.test()` has been extended to handle one-sample and paired tests.
- The `palette()` function has a new default set of colours (which are less saturated and have better accessibility properties). There are also some new built-in palettes, which are listed by the new `palette.pals()` function. These include the old default palette under the name `"R3"`. Finally, the new `palette.colors()` function allows a subset of colours to be selected from any of the built-in palettes.
- `n2mfrow()` gains an option `asp = 1` to specify the aspect ratio, fulfilling the wish and extending the proposal of Michael Chirico in PR#17648.
- For `head(x, n)` and `tail()` the default and other S3 methods notably for *vector* `n`, e.g. to get a "corner" of a matrix, has been extended to `array`'s of higher dimension thanks to the patch proposal by Gabe Becker in PR#17652. Consequently, optional argument `addrownums` is deprecated and replaced by the (more general) argument `keepnums`. An invalid second argument `n` now leads to typically more easily readable error messages.
- New function `.class2()` provides the full character vector of class names used for S3 method dispatch.
- Printing `methods(..)` now uses a new `format()` method.
- `sort.list(x)` now works for non-atomic objects `x` and `method = "auto"` (the default) or `"radix"` in cases `order(x)` works, typically via a `xtfrm()` method.
- Where they are available, `writeBin()` allows long vectors.
- New function `deparse1()` produces one string, wrapping `deparse()`, to be used typically in `deparse1(substitute(*))`, e.g., to fix PR#17671.
- `wilcox.test()` enhancements: In the (non-paired) two-sample case, `Inf` values are treated as very large for robustness consistency. If exact computations are used, the result now has `"exact"` in the `method` element of its return value. New arguments `tol.root` and `digits.rank` where the latter may be used for stability to treat very close numbers as ties.
- `readBin()` and `writeBin()` now report an error for an invalid `endian` value. The affected code needs to be fixed with care as the old undocumented behavior was to swap endian-ness in such cases.

- `sequence()` is now an S3 generic with an internally implemented default method, and gains arguments to generate more complex sequences. Based on code from the **S4Vectors** Bioconductor package and the advice of Hervé Pagès.
- `print()`'s default method and many other methods (by calling the default eventually and passing `...`) now make use of a new optional `width` argument, avoiding the need for the user to set and reset `options("width")`.
- `memDecompress()` supports the RFC 1952 format (e.g. in-memory copies of `gzip`-compressed files) as well as RFC 1950.
- `memCompress()` and `memDecompress()` support long raw vectors for types `"gzip"` and `"zx"`.
- `sweep()` and `slice.index()` can now use names of dimnames for their `MARGIN` argument (`apply` has had this for almost a decade).
- New function `proportions()` and `marginSums()`. These should replace the unfortunately named `prop.table()` and `margin.table()`. They are drop-in replacements, but also add named-margin functionality. The old function names are retained as aliases for back-compatibility.
- Functions `rbinom()`, `rgeom()`, `rhyper()`, `rpois()`, `rnbinom()`, `rsignrank()` and `rwilcox()` which have returned integer since R 3.0.0 and hence `NA` when the numbers would have been outside the integer range, now return double vectors (without NAs, typically) in these cases.
- `matplot(x,y)` (and hence `matlines()` and `matpoints()`) now call the corresponding methods of `plot()` and `lines()`, e.g, when `x` is a `"Date"` or `"POSIXct"` object; prompted by Spencer Graves' suggestion.
- `stopifnot()` now allows customizing error messages via argument names, thanks to a patch proposal by Neal Fultz in PR#17688.
- `unlink()` gains a new argument `expand` to disable wildcard and tilde expansion. Elements of `x` of value `"~"` are now ignored.
- `mle()` in the `stats4` package has had its interface extended so that arguments to the negative log-likelihood function can be one or more vectors, with similar conventions applying to bounds, start values, and parameter values to be kept fixed. This required a minor extension to class `"mle"`, so saved objects from earlier versions may need to be recomputed.
- The default for `pdf()` is now `useDingbats = FALSE`.
- The default fill colour for `hist()` and `boxplot()` is now `col = "lightgray"`.
- The default order of the levels on the y-axis for `spineplot()` and `cdplot()` has been reversed.
- If the `R_ALWAYS_INSTALL_TESTS` environment variable is set to a true value, `R CMD INSTALL` behaves as if the '`--install-tests`' option is always specified. Thanks to Reinhold Koch for the suggestion.
- New function `R_user_dir()` in package **tools** suggests paths appropriate for storing R-related user-specific data, configuration and cache files.
- `capabilities()` gains a new logical option `Xchk` to avoid warnings about X11-related capabilities.
- The internal implementation of **grid** units has changed, but the only visible effects at user-level should be
  - a slightly different print format for some units (especially unit arithmetic),
  - faster performance (for unit operations) and
  - two new functions `unitType()` and `unit.psum()`.

  Based on code contributed by Thomas Lin Pedersen.

- When internal dispatch for `rep.int()` and `rep_len()` fails, there is an attempt to dispatch on the equivalent call to `rep()`.
- Object `.Machine` now contains new `longdouble.*` entries (when R uses long doubles internally).
- `news()` has been enhanced to cover the news on R 3.x and 2.x.
- For consistency, `N <- NULL; N[[1]] <- val` now turns `N` into a `list` also when `val)` has length one. This enables `dimnames(r1)[[1]] <- "R1"` for a 1-row matrix `r1`, fixing PR#17719 reported by Serguei Sokol.
- `deparse(..)`, `dump(..)`, and `dput(x, control = "all")` now include control option `"digits17"` which typically ensures 1:1 invertibility. New option `control = "exact"` ensures numeric exact invertibility via `"hexNumeric"`.
- When loading data sets via `read.table()`, `data()` now uses '`LC_COLLATE=C`' to ensure locale-independent results for possible string-to-factor conversions.
- A server socket connection, a new connection type representing a listening server socket, is created via `serverSocket()` and can accept multiple socket connections via `socketAccept()`.
- New function `socketTimeout()` changes the connection timeout of a socket connection.
- The time needed to start a homogeneous '`PSOCK`' cluster on '`localhost`' with many nodes has been significantly reduced (package **parallel**).
- New `globalCallingHandlers()` function to establish global condition handlers. This allows registering default handlers for specific condition classes. Developed in collaboration with Lionel Henry.
- New function `tryInvokeRestart()` to invoke a specified restart if one is available and return without signaling an error if no such restart is found. Contributed by Lionel Henry in PR#17598.
- `str(x)` now shows the length of `attributes` in some cases for a data frame `x`.
- `Rprof()` gains a new argument `filter.callframes` to request that intervening call frames due to lazy evaluation or explicit `eval()` calls be omitted from the recorded profile data. Contributed by Lionel Henry in PR#17595.
- The handling of `${FOO-bar}` and `${FOO:-bar}` in '`Renviron`' files now follows POSIX shells (at least on a Unix-alike), so the first treats empty environment variables as set and the second does not. Previously both ignored empty variables. There are several uses of the first form in '`etc/Renviron`'.
- New `classes` argument for `suppressWarnings()` and `suppressMessages()` to selectively suppress only warnings or messages that inherit from particular classes. Based on patch from Lionel Henry submitted with PR#17619.
- New function `activeBindingFunction()` retrieves the function of an active binding.
- New `"cairoFT"` and `"pango"` components in the output of `grSoftVersion()`.
- New argument `symbolfamily` in cairo-based graphics devices and new function `cairoSymbolFont()` that can be used to provide the value for that argument.

Windows:

- `Rterm` now works also when invoked from MSYS2 terminals. Line editing is possible when command `winpty` is installed.
- `normalizePath()` now resolves symbolic links and normalizes case of long names of path elements in case-insensitive folders (PR#17165).
- `md5sum()` supports UTF-8 file names with characters that cannot be translated to the native encoding (PR#17633).

- `Rterm` gains a new option '`--workspace`' to specify the workspace to be restored. This allows equals to be part of the name when opening *via* Windows file associations (reported by Christian Asseburg).
- `Rterm` now accepts `ALT+xxx` sequences also with NumLock on. Tilde can be pasted with an Italian keyboard (PR#17679).
- R falls back to copying when junction creation fails during package checking (patch from Duncan Murdoch).

DEPRECATED AND DEFUNCT:

- Make macro '`F77_VISIBILITY`' has been removed and replaced by '`F_VISIBILITY`'.
- Make macros '`F77`', '`FCPIFCPLAGS`' and '`SHLIB_OPENMP_FCFLAGS`' have been removed and replaced by '`FC`', '`FPICFLAGS`' and '`SHLIB_OPENMP_FFLAGS`' respectively. (Most `make` programs will set '`F77`' to the value of '`FC`', which is set for package compilation. But portable code should not rely on this.)
- The deprecated support for specifying C++98 for package installation has been removed.
- `R CMD config` no longer knows about the unused settings '`F77`' and '`FCPIFCPLAGS`', nor '`CXX98`' and similar.
- Either PCRE2 or PCRE1 >= 8.32 (Nov 2012) is required: the deprecated provision for 8.20–8.31 has been removed.
- Defunct functions `mem.limits()`, `.readRDS()`, `.saveRDS()`, `.find.package()`, and `.path.package()` from package **base** and `allGenerics()`, `getAccess()`, `getAllMethods()`, `getClassName()`, `getClassPackage()`, `getExtends()`, `getProperties()`, `getPrototype()`, `getSubclasses()`, `getVirtual()`, `mlistMetaName()`, `removeMethodsObject()`, `seemsS4Object()`, `traceOff()`, and `traceOn()` from **methods** have been removed.

C-LEVEL FACILITIES:

- `installChar` is now remapped in '`Rinternals.h`' to `installTrChar`, of which it has been a wrapper since R 3.6.0. Neither are part of the API, but packages using `installChar` can replace it if they depend on '`R >= 3.6.2`'.
- Header '`R_ext/Print.h`' defines '`R_USE_C99_IN_CXX`' and hence exposes `Rvprintf` and `REvprintf` if used with a C++11 (or later) compiler.
- There are new Fortran subroutines `dblepr1`, `realpr1` and `intpr1` to print a scalar variable (`gfortran` 10 enforces the distinction between scalars and length-one arrays). Also `labelpr` to print just a label.
- `R_withCallingErrorHandler` is now available for establishing a calling handler in C code for conditions inheriting from class `error`.

INSTALLATION on a UNIX-ALIKE:

- User-set '`DEFS`' (e.g., in '`config.site`') is now used for compiling packages (including base packages).
- There is a new variant option '`--enable-lto=check`' for checking consistency of BLAS/LAPACK/LINPACK calls — see 'Writing R Extensions'.
- A C++ compiler default is set only if the C++11 standard is supported: it no longer falls back to C++98.
- PCRE2 is used if available. To make use of PCRE1 if PCRE2 is unavailable, configure with option '`--with-pcre1`'.
- The minimum required version of `libcurl` is now 7.28.0 (Oct 2012).
- New make target `distcheck` checks

- R can be rebuilt from the tarball created by `make dist`,
- the build from the tarball passes `make check-all`,
- the build installs and uninstalls,
- the source files are properly cleaned by `make distclean`.

UTILITIES:

- `R --help` now mentions the option `--no-echo` (renamed from `--slave`) and its previously undocumented short form `-s`.
- `R CMD check` now optionally checks `configure` and `cleanup` scripts for non-Bourne-shell code ('bashisms').
- `R CMD check --as-cran` now runs `\donttest` examples (which are run by `example()`) instead of instructing the tester to do so. This can be temporarily circumvented during development by setting environment variable `_R_CHECK_DONTTEST_EXAMPLES_` to a false value.

PACKAGE INSTALLATION:

- There is the beginnings of support for the recently approved C++20 standard, specified analogously to C++14 and C++17. There is currently only limited support for this in compilers, with flags such as '`-std=c++20`' and '`-std=c++2a`'. For the time being the `configure` test is of accepting one of these flags and compiling C++17 code.

BUG FIXES:

- `formula(x)` with `length(x) > 1` character vectors, is deprecated now. Such use has been rare, and has 'worked' as expected in some cases only. In other cases, wrong `x` have silently been truncated, not detecting previous errors.
- Long-standing issue where the X11 device could lose events shortly after startup has been addressed (PR#16702).
- The `data.frame` method for `rbind()` no longer drops `<NA>` levels from factor columns by default (PR#17562).
- `available.packages()` and hence `install.packages()` now pass their `...` argument to `download.file()`, fulfilling the wish of PR#17532; subsequently, `available.packages()` gets new argument `quiet`, solving PR#17573.
- `stopifnot()` gets new argument `exprObject` to allow an R object of class `expression` (or other 'language') to work more consistently, thanks to suggestions by Suharto Anggono.
- `conformMethod()` now works correctly in cases containing a "`&&` logic" bug, reported by Henrik Bengtsson. It now creates methods with `"missing"` entries in the signature. Consequently, `rematchDefinition()` is amended to use appropriate `.local()` calls with named arguments where needed.
- `format.default(*, scientific = FALSE)` now corresponds to a practically most extreme `options(scipen = n)` setting rather than arbitrary `n = 100`.
- `format(as.symbol("foo"))` now works (returning `"foo"`).
- `postscript(.., title = *)` now signals an error when the title string contains a character which would produce corrupt PostScript, thanks to PR#17607 by Daisuko Ogawa.
- Certain `Ops` (notably comparison such as `==`) now also work for 0-length data frames, after reports by Hilmar Berger.
- `methods(class = class(glm(..)))` now warns more usefully and only once.
- `write.dcf()` no longer mangles field names (PR#17589).

- Primitive replacement functions no longer mutate a referenced first argument when used outside of a complex assignment context.
- A better error message for `contour(*, levels = Inf)`.
- The return value of `contourLines()` is no longer `invisible()`.
- The Fortran code for calculating the `coefficients` component in `lm.influence()` was very inefficient. It has (for now) been replaced with much faster R code (PR#17624).
- `cm.colors(n)` *etc* no longer append the code for `alpha = 1`, `"FF"`, to all colors. Hence all eight `*.colors()` functions and `rainbow()` behave consistently and have the same non-explicit default (PR#17659).
- `dnorm` had a problematic corner case with `sd == -Inf` or negative `sd` which was not flagged as an error in all cases. Thanks to Stephen D. Weigand for reporting and Wang Jiefei for analyzing this; similar change has been made in `dlnorm()`.
- The optional `iter.smooth` argument of `plot.lm()`, (the `plot()` method for `lm` and `glm` fits) now defaults to `0` for all `glm` fits. Especially for binary observations with high or low fitted probabilities, this effectively deleted all observations of 1 or 0. Also, the type of residuals used in the `glm` case has been switched to `"pearson"` since deviance residuals do not in general have approximately zero mean.
- In `plot.lm`, Cook's distance was computed from unweighted residuals, leading to inconsistencies. Replaced with usual weighted version. (PR#16056)
- Time-series `ts(*, start, end, frequency)` with fractional `frequency` are supported more consistently; thanks to a report from Johann Kleinbub and analysis and patch by Duncan Murdoch in PR#17669.
- In case of errors `mcmapply()` now preserves attributes of returned `"try-error"` objects and avoids simplification, overriding `SIMPLIFY` to `FALSE`. (PR#17653)
- `as.difftime()` gets new optional `tz = "UTC"` argument which should fix behaviour during daylight-savings-changeover days, fixing PR#16764, thanks to proposals and analysis by Johannes Ranke and Kirill Müller.
- `round()` does a better job of rounding *"to nearest"* by *measuring* and *"to even"*; thanks to a careful algorithm originally prompted by the report from Adam Wheeler and then others, in PR#17668.
  `round(x, dig)` for *negative* digits is much more rational now, notably for large $|dig|$.
- Inheritance information on S4 classes is maintained more consistently, particularly in the case of class unions (in part due to PR#17596 and a report from Ezra Tucker).
- `is()` behaves more robustly when its argument `class2` is a `classRepresentation` object.
- The warning message when attempting to export an nonexistent class is now more readable; thanks to Thierry Onkelinx for recognizing the problem.
- `choose()` misbehaved in corner cases where it switched `n - k` for `k` and `n` was only *nearly* integer (report from Erik Scott Wright).
- `mle()` in the `stats4` package had problems combining use of box constraints and fixed starting values (in particular, confidence intervals were affected).
- Operator `?` now has lower precedence than `=` to work as documented, so `=` behaves like `<-` in help expressions (PR#16710).
- `smoothEnds(x)` now returns `integer` type in *both* cases when `x` is `integer`, thanks to a report and proposal by Bill Dunlap PR#17693.
- The **methods** package does a better job of tracking inheritance relationships across packages.
- `norm(diag(c(1, NA)), "2")` now works.

- `subset()` had problems with 0-col dataframes (reported by Bill Dunlap, PR#17721).
- Several cases of integer overflow detected by the 'undefined behaviour sanitizer' of `clang` 10 have been circumvented. One in `rhyper()` may change the generated value for large input values.
- `dotchart()` now places the y-axis label (`ylab`) much better, not overplotting labels, thanks to a report and suggestion by Alexey Shipunov.
- A rare C-level array overflow in `chull()` has been worked around.
- Some invalid specifications of the day-of-the-year (*via* `%j`, e.g. day 366 in 2017) or week plus day-of-the-week are now detected by `strptime()`. They now return `NA` but give a warning as they may have given random results or corrupted memory in earlier versions of R.
- `socketConnection(server = FALSE)` now respects the connection timeout also on Linux.
- `socketConnection(server = FALSE)` no longer leaks a connection that is available right away without waiting (e.g. on '`localhost`').
- Socket connections are now robust against spurious readability and spurious availability of an incoming connection.
- `blocking = FALSE` is now respected also on the server side of a socket connection, allowing non-blocking read operations.
- `anova.glm()` and `anova.glmlist()` computed incorrect score (Rao) tests in no-intercept cases. (André Gillibert, PR#17735)
- `summaryRprof()` now should work correctly for the `Rprof(*, memory.profiling=TRUE)` case with small chunk size (and `"tseries"` or similar) thanks to a patch proposal by Benjamin Tyner, in PR#15886.
- `xgettext()` ignores strings passed to `ngettext()`, since the latter is handled by `xngettext()`. Thanks to Daniele Medri for the report and all the recent work he has done on the Italian translations.
- `data(package = "P")` for `P` in **base** and **stats** no longer reports the data sets from package **datasets** (which it did for back compatibility for 16 years), fixing PR#17730.
- `x[[Inf]]` (returning `NULL`) no longer leads to undefined behavior, thanks to a report by Kirill Müller in PR#17756. Further, `x[[-Inf]]` and `x[[-n]]` now give more helpful error messages.
- `Gamma()` family sometimes had trouble storing link name PR#15891

BUG FIXES (Windows):

- `Sys.glob()` now supports all characters from the Unicode Basic Multilingual Plane, no longer corrupting some (less commonly used) characters (PR#17638).
- `Rterm` now correctly displays multi-byte-coded characters representable in the current native encoding (at least on Windows 10 they were sometimes omitted, PR#17632).
- `scan()` issues with UTF-8 data when running in a DBCS locale have been resolved (PR#16520, PR#16584).
- `Rterm` now accepts enhanced/arrow keys also with ConPTY.
- R can now be started *via* the launcher icon in a user documents directory whose path is not representable in the system encoding.
- `socketConnection(server = FALSE)` now returns instantly also on Windows when connection failure is signalled.
- Problems with UTF-16 surrogate pairs have been fixed in several functions, including `tolower()` and `toupper()` (PR#17645).

CHANGES in previous versions

- Older news can be found in text format in files 'NEWS.0', 'NEWS.1', 'NEWS.2' and 'NEWS.3' in the 'doc' directory. News in HTML format for R versions 3.x and from 2.10.0 to 2.15.3 are available at 'doc/html/NEWS.3.html' and 'doc/html/NEWS.2.html'.