

Bibcheck

Version 1.2 (2023/02/01)

Bibcheck is a Lua-based script that checks each \backslash bibitem of a \LaTeX file (.tex) against MathSciNet and zbMATH, and writes all checked entries into a Bib \TeX file (.bib).

Bibcheck is available at <https://github.com/ems-press/bibcheck>.

For bug reports, comments and suggestions contact [Simon Winter](#).

Contents

1	Installation	1
1.1	Install Lua v5.1 (or higher) on Windows	1
1.2	Install WGet v1.14 (or higher) on Windows	1
1.3	Install Lua and WGet on Mac/Linux	2
1.4	Install Bibcheck	2
2	How to use Bibcheck	2
2.1	Case 1: \backslash bibitem's	2
2.2	Case 2: Bib \TeX	3
2.3	Case 3: amsrefs	3
2.4	Use a .bat file on Windows	4
3	What to pay attention to	4
4	The config file	5
5	Changelog	5

1 Installation

1.1 Install Lua v5.1 (or higher) on Windows

1. Download LuaForWindows_v5.1.5-52.exe (27.8 MB) at <https://github.com/rjpcomputing/luaforwindows/releases/tag/v5.1.5-52>
2. Run the .exe file and always click 'accept/next'.

1.2 Install WGet v1.14 (or higher) on Windows

1. Download from <https://eternallybored.org/misc/wget/> the .exe (!) file of Version 1.21 (32-bit or 64-bit).
2. Copy wget.exe to e.g. C:\Program Files (x86)\wget-1.21.1-1-win64\ or any other folder.
3. Add WGet to the Windows PATH:
 - (a) Open the Start Search, type in 'env' and choose 'Edit the system environment variables'.
 - (b) Click the 'Environment Variables' button.
 - (c) Under 'System Variables' find the row with 'Path' in the first column and click edit.
 - (d) Click 'New' and type in the new path, e.g. C:\Program Files (x86)\wget-1.21.1-1-win64
 - (e) Dismiss all of the dialogs by choosing OK. Your changes are saved.

4. To check whether the installation was successful, open a command terminal (by typing 'cmd' in the search menu) and type

```
wget -qO-  
"https://mathscinet.ams.org/mathscinet-mref?dataType=bibtex&ref=J. Whitehead,  
On 2-spheres in 3-manifolds. Bull. Amer. Math. Soc. 64 (1958), 161--166"
```

or

```
wget -qO-  
"https://zbmath.org/citationmatching/match?f=latex&q=J. Whitehead,  
On 2-spheres in 3-manifolds. Bull. Amer. Math. Soc. 64 (1958), 161--166"
```

in both cases without any line break. The return should be MR 103473 resp. Zbl0084.19103.

If this doesn't work, try again with `wget --no-check-certificate -qO- ...` If only this works, see §1.4.4.

1.3 Install Lua and WGet on Mac/Linux

1. We recommend to use the package manager [Homebrew](#).
2. See [Method #1](#) on how to install Homebrew and WGet.
3. After installing Lua, the two Lua libraries 'LuaFileSystem' and 'Penlight' might be missing. Then run:

```
brew install luarocks  
luarocks install luafilesystem  
luarocks install penlight
```

1.4 Install Bibcheck

1. Create a new folder for Bibcheck. *The whole path must not contain any spaces!*
2. Copy all three .lua files into the Bibcheck folder: bibcheck.lua, config.lua, functions.lua
3. Download dkjson.lua (v2.5) from <http://dkolf.de/src/dkjson-lua.fsl/home> and copy the file into the Bibcheck folder.¹
4. It could be (e.g. under Windows 11) that you need to replace 'wget -qO-' in line 112 and/or 123 of bibcheck.lua by 'wget --no-check-certificate -qO-'; see the test described in §1.2.4.²

2 How to use Bibcheck

2.1 Case 1: \bibitem's

If your .tex file contains a thebibliography environment³, proceed as follows:

1. Open the command terminal, go to the paper's directory and write
lua PATH\bibcheck.lua MAINFILE.tex BSTFILENAME
e.g. on Windows:
lua C:\tools\bibcheck\bibcheck.lua main.tex ems

Note: *The whole path must not contain any spaces!*

2. The script will run for a few seconds. What happens is the following: Every \bibitem in main.tex is compared with MathSciNet.
 - If there is a match, this **match** is written into a .bib file.
 - If there is no match, the **original entry** is written into the .bib file.

¹Mac users could probably use luarocks to install dkjson.lua; see §1.3.3.

²This will be improved in the next version of Bibcheck.

³This is the default case for Bibcheck.

3. Next, an automatic run of latex and bibtex creates a .bbl file. This .bbl file is pasted into the original .tex file main.tex., which is then renamed main_bibchecked.tex.
4. Now you can start editing the new ('bibchecked') .tex file. For this, *you must check each \bibitem*:
 - If there was a **MathSciNet match**, the original entry appears as a comment (%) below the match. It is important to compare the original entry and the match because sometimes MathSciNet delivers a wrong result. However, a quick comparison of e.g. the page range provides clarity. Or, if the original entry has an MR number, you can compare it with the MR number of the match.
 - If there was **no MathSciNet match**, the entry is now at the beginning of the bibliography. This means, you have to (a) sort this entry and (b) format it by hand. This is a disadvantage of the tool; but it happens to only a few entries.
5. In addition, each \bibitem is compared with zbMATH and (if there was a match) extended by the respective number. But again: You must compare the original entry and the zbMATH match (which also appears below the match as a comment) to ensure that both describe the same publication.
6. That's it!

2.2 Case 2: BibTeX

If there is a .bib file, say bibsource.bib, proceed as follows:

1. In the main .tex file write:


```
\bibliographystyle{ems}
\bibliography{bibsource}
```
2. Run bibtex to create bibsource.bbl.
3. Copy the content of bibsource.bbl into the .tex file.
4. Proceed as in Case 1.⁴

2.3 Case 3: amsrefs

If the bibliography is prepared using the amsrefs package, proceed as follows:

1. Copy all \bib entries into a new .ltx file, say source.ltx.
2. Create a temporary .tex file, say temp.tex, consisting of the following lines:⁵

```
\documentclass{article}
\usepackage{amsrefs,ltx2bib}
\begin{document}
\begin{bibdiv}
\begin{biblist}
\bibselect*{source}
\end{biblist}
\end{bibdiv}
\writebib{source}{bibsource}
\end{document}
```
3. Compile temp.tex. This will create a .bib file bibsource.bib containing all \bib entries from source.ltx.

Warning: The .bib file is not perfect: wrong brackets around author names, wrong structure of @InCollection entries, capitalization of proper names is lost (due to missing brackets), and maybe more. However, these conversion errors should have little effect on the hit rate in MathSciNet and zbMATH.
4. In the *original* .tex file (say, main.tex), replace the bibliography by


```
\bibliographystyle{ems}
\bibliography{bibsource}
```

⁴Isn't this procedure totally cumbersome? Yes, it is. Maybe someday a new version of Bibcheck will make it better.

⁵The reason we create a new .tex file is that there is a conflict with the \title command: if the .tex file contains \title, then all article and book titles disappear when converting ltx to bib.

5. Run `bibtex`. But before you do this, you must remove `\usepackage{amsrefs}` (otherwise the `.bib` file will be in `amsrefs` format).
6. Copy the content of `main.bbl` into the `.tex` file.
7. Now you can finally use Bibcheck as in Case 1.
8. **Warning:** What Bibcheck considers to be the “original entry” is of course only the entry from the defective `.bib` file created in step 2.3.3. Therefore, in cases of doubt, the original *manuscript* should be consulted for comparison.⁶

2.4 Use a .bat file on Windows

Instead of using the command terminal, you can create and use a batch file (`.bat`):

1. Create, using any text editor, a file `bibcheck.bat` containing the following four lines:

```
@echo off
chcp 65001
lua "C :...\ bibcheck.lua" %~f1 ems
pause
```

Here `C :...\ bibcheck.lua` is the full path of `bibcheck.lua`. Keep the quotation marks!

2. Create a desktop shortcut of that `.bat` file.
3. Now drag and drop your `.tex` file onto the desktop shortcut. Easy!

3 What to pay attention to

1. Before you bibcheck a `.tex` file, you should
 - replace each `\bysame` by the respective authors.
2. Be aware of **mismatches**. They mainly appear when
 - the original entry is a preprint (e.g. arXiv) and MathSciNet returns a journal article with the same title;
 - the original entry is the original version (e.g. in Russian) and MathSciNet returns the English translation, or the other way around;
 - the original entry and the MathSciNet match have a very similar title, maybe even the same author, but are different.
3. The bibliography in `main_bibchecked.tex` is based on MathSciNet. Here are some known problems with MathSciNet:
 - Book **series** are often not abbreviated; so you must abbreviate them by hand following <https://www.siam.org/Portals/0/Journal%20PDFs/serials.pdf>.
 - Write ‘Springer’ instead of ‘Springer-Verlag’, ‘Springer Verlag’, etc. Same with ‘Birkhäuser’: remove ‘Verlag’.
 - Some **Astérisque** papers are of entry type `@incollection` containing a field journal:


```
@incollection {Shelstad1989,
    AUTHOR = {Shelstad, D.},
    TITLE = {A formula for regular unipotent germs},
    NOTE = {Orbites unipotentes et repr{\{e\}}sentations, II},
    JOURNAL = {Ast{\{e\}}risque},
    FJOURNAL = {Ast{\{e\}}risque},
    NUMBER = {171-172},
    YEAR = {1989},
    PAGES = {275--277},
    ISSN = {0303-1179},
```

⁶Instead of using this `lbt2bib` approach, it might make more sense to use the `\bib` entries from `amsrefs` as the input for Bibcheck (i.e. without the detour via `BIBTEX`). One probably only has to remove the fields names (`author=`, `title=`, etc.) and replace `\bib` with `\bibitem`. Maybe Bibcheck will be able to do that in one of the next versions.

```
MRCLASS = {22E35 (11F70 11R39 11S37)},
MRNUMBER = {1021506},
MRREVIEWER = {Joe Repka},
}
```

Officially, `journal` is not an allowed field for `@incollection`. Thus, `.bst` files (such as `ems.bst`) will not print `journal`, they rather expect a field `booktitle`. In this case, the name ‘Astérisque’ has to be restored by hand.

4. The hit rate with `zbMATH` is not as high as with `MathSciNet`. This means: Not all publications listed in `zbMATH` are found. In this case, a manual search often helps: <https://zbmath.org/citationmatching/>. Here, the hit probability increases if you only enter *parts* of the `\bibitem`.
5. MR numbers are returned without leading zero, e.g. `\MR{0553218}` \rightarrow `\MR{553218}`.

4 The config file

1. When running `bibcheck`, you may skip the last argument (`BSTFILENAME`). This will activate the default bibliography style as defined in `config.lua`:

```
M.bibstyle = 'ems'
```

2. The temporary `.bib` file contains the MR and `zbMATH` numbers:

```
zblnumber = {1460.22007}, mrnumber = {4028458}
```

If your `.bst` file knows how to deal with such entries (e.g. `ems.bst` does), the final `\bibitem` will contain

```
\Zbl{1460.22007} \MR{4028458}
```

The `.tex/.cls/.sty` file, in turn, must have a definition for the `\Zbl` and `\MR` commands.

This means, if you don’t need any `zbMATH` numbers, you may either

- use a `.bst` file that ignores the field `zblnumber`, or
- define `\newcommand\Zbl[1]{}{}`, or
- change the respective line in `config.lua` to

```
M.printZbl = false
```

The latter option disables the comparison with `zbMATH` and makes `Bibcheck` run much faster!

3. If you don’t need DOIs,

- use a `.bst` file that ignores the field `doi`, *and*
- change the respective line in `config.lua` to

```
M.checkCrossref = false
```

4. If you want to keep the temporary `.bib` file, remove the respective entry from the list of removed files in `config.lua`:

```
M.remove_files = { -- '.bib',
  '.bbl', '.aux', '.log', '.dvi', '.blg', }
```

5 Changelog

- Version 1.1
 - `split_at_bibitem`: removes blank lines inside any `\bibitem`.
 - `make_bib` and `space_warning`: removes spaces in `\bibitem` labels and adds a warning in `main_bibchecked.tex`.
 - `make_bib` and `undress`: curly braces and tex commands of the form `\XYZ` removed before the `\bibitem` is sent to the `zbMATH` Citation Matcher.

- zbl_ID: Bug fixed.
 - escape_percents: Function added; otherwise strings such as \%3 (e.g. in URLs) wouldn't work in Lua 5.1.
- Version 1.2
 - make_bib and undress: further simplify each \bibitem before it is sent to the zbMATH Citation Matcher. This way we try to increase the zbMATH hit rate.
 - make_bib: DOIs added. If the MathSciNet match doesn't contain any DOI, the original entry is sent to the Crossref API.