The XyMTeX system for publishing interdisciplinary chemistry/mathematics books

Shinsaku Fujita

1 XyMTeX Version 5.01

I have recently released XyMTeX Version 5.01 for drawing chemical structural formulas, where its zip file (`xymtx501.zip`) is available from my personal homepage (`http://xymtex.com/`). I have more recently uploaded this version to the CTAN archives.

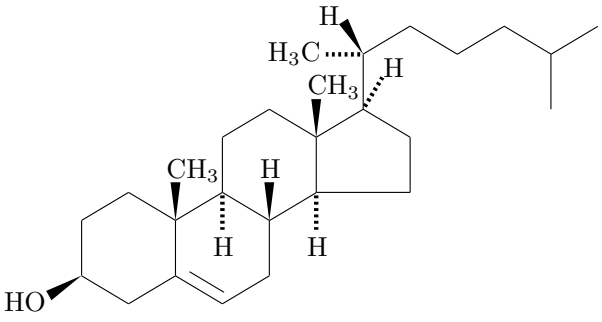
The XyMTeX system supports three modes for drawing:

1. the LaTeX-compatible mode, which is based on the LaTeX picture environment along with the epic package,
2. the PostScript-compatible mode, which is based on the PSTricks package, and
3. the PDF-compatible mode, which is based on the pgf/TikZ package.

The three modes can be switched by loading the xymtex, xymtexps, or xymtexpdf package with the `\usepackage` command. If structural formulas of high quality are necessary, the latter two modes should be selected. A typical template for switching the three modes is shown below:

```
\documentclass{article}
%\usepackage{xymtex}  %LaTeX mode
%\usepackage{xymtexps}%PostScript mode
\usepackage{xymtexpdf}%PDF mode
\usepackage{graphicx}
\begin{document}
\cholestane[e]{3B==HO}%XyMTeX command
\end{document}
```

The XyMTeX command `\cholestane` with the arguments `[e]` and `{3B==HO}` generates the chemical structural formula of cholest-5-en-3β-ol as follows:



Because PDF is now a default standard for exchanging digital documents, it is usually highly desirable to convert DVI files obtained by the PostScript-compatible mode or the PDF-compatible mode to PDF files. To obtain a PDF file of printing quality, the following routes are typical:

1. PostScript-compatible mode:
   As the more classical process, a `dvi` file produced by the PostScript-compatible mode is converted into a `ps` file. The resulting `ps` file is in turn converted into a `pdf` file.
   $$\text{tex} \overset{\text{LaTeX}}{\longrightarrow} \text{dvi} \overset{\text{dvips}}{\longrightarrow} \text{ps} \overset{\text{distiller}}{\longrightarrow} \text{pdf}$$
2. PDF-compatible mode:
   - Because PDF technology has become predominant over the PostScript technology, a `dvi` file produced by the PDF-compatible mode is directly converted into a `pdf` file by using the `dvipdfmx` converter.
     $$\text{tex} \overset{\text{LaTeX}}{\longrightarrow} \text{dvi} \overset{\text{dvipdfmx}}{\longrightarrow} \text{pdf}$$
   - The PDF-compatible mode can take an optional argument `pdftex` as follows:
     `\usepackage[pdftex]{xymtexpdf}`
     Thereby, a `tex` file is directly converted into a `pdf` file by using the `pdflatex` engine:
     $$\text{tex} \overset{\text{pdflatex}}{\longrightarrow} \text{pdf}$$

It should be emphasized that common code written for the XyMTeX system can be used in any of the routes itemized above.

2 Techniques for drawing complicated structural formulas

XyMTeX commands are equipped with facilities for drawing complex structures, i.e., the substitution technique for attaching substituents, the addition technique for drawing fused rings, and the replacement technique for drawing spiro rings. The detailed documentation of the XyMTeX system [1] is available from my homepage located at `http://xymtex.com/`.

2.1 The substitution technique

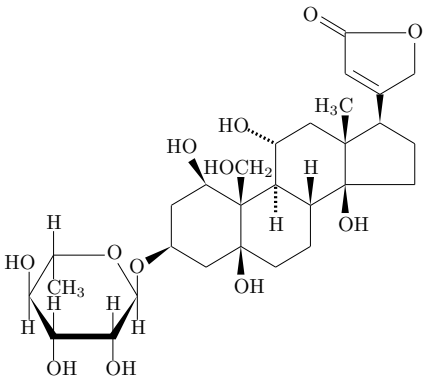
The XyMTeX system supports the substitution technique, which is based on (yl)-functions for linking complicated substituents. An intervening divalent unit can be inserted by using a command `\ryl` or `\lyl`.

For example, the structural formula of *g*-strophanthin (ouabain) as a poisonous cardiac glycoside is drawn by the code [1]:

```
\begin{XyMcompd}(2000,1850)(-550,-300){}{}
\steroid{1SB==\lmoiety{HO};5B==OH;8B==H;%
9A==H;{11}A==HO;{10}B==\llap{HO}CH$_{2}$;%
{14}B==OH;{13}B==\lmoiety{H$_{3}$C};%
{17}B==\fiveheterov[e]{3==O}%
{4D==O;1==(yl)};3B==\lyl(3==O){8==%
```

```
\pyranosew{1==(yl);1Sa==H;2Sb==H;2Sa==OH;%
3Sb==H;3Sa==OH;4Sb==HO;%
4Sa==H;5Sb==H;5Sa==CH$_{3}$}}}
\end{XyMcompd}
```

In this code, a steroid skeleton (due to the `\steroid` command) is substituted by a five-membered heterocycle (due to a (yl)-function in the `\fiveheterov` command) and by a pyranose moiety (due to a (yl)-function in the `\pyranosew` command and a further use of the `\lyl` command). The `XyMcompd` environment secures a drawing area for the structure to be drawn. This code typesets the following structural formula:

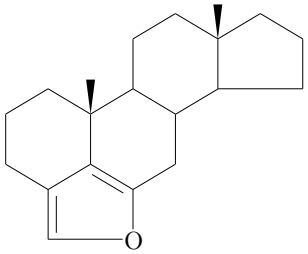


2.2 The addition technique

The XƎMTEX system supports the addition technique, where the attachment mechanism permits a given fusing unit to be attached onto an edge of a parent skeleton.

For example, furo[40,30,20:4,5,6]androstane as a fused steroid is drawn by the addition technique, where the `\fivefusevi` command for drawing a 5-membered fusing unit is declared in the bond list of the `\steroid` command for drawing a steroid skeleton:

```
\steroid
[{c{\fivefusevi[ad]{3==O}{}{e}[a]}}]
{{10}B==\null;{13}B==\null}
```

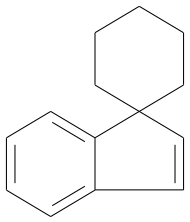


2.3 The replacement technique

The XƎMTEX system supports the replacement technique, where a spiro unit is drawn on the basis of a (yl)-function and attached to a vertex of a parent skeleton.

For example, spiro[cyclohexane-1,1′-indene] is drawn by the replacement technique, where a six-membered spiro unit is produced by declaring a (yl)-function in the `\cyclohexanev` command:

```
\begin{XyMcompd}(600,800)(250,250){}{}
\nonaheterovi[bdfh]%
{1s==\cyclohexanev{4==(yl)}}{}
\end{XyMcompd}
```



3 Interdisciplinary chemistry/mathematics books

The development of the XƎMTEX system largely reflects the personal history of my research aiming at the integration of chemistry and mathematics, e.g., the concept of imaginary transition structures (ITSs) [2], the USCI (unit-subduced-cycle-index) approach [3, 4], the concept of stereoisograms [5, 6], the proligand method [7], and the concept of mandalas [8].

3.1 Manual drawing without using the XƎMTEX system

In 1991, I published an interdisciplinary monograph on the combinatorial enumeration of chemical compounds as three-dimensional structures (the USCI approach) [9]. This book contains many structural formulas of organic compounds along with mathematical equations because of its interdisciplinary nature. Such mathematical equations were successfully typeset by means of the original programs of the (LA)TEX system. However, the structural formulas contained in this book were drawn manually and pasted on the camera-ready manuscript, because (LA)TEX at that time had no reliable utility for drawing structural formulas, and commercially available systems such as ChemDraw were too expensive to be used for personal purposes.

The concept of imaginary transition structures (ITSs), which serve as computer-oriented representations of organic reactions, was developed mainly during the 1980s. In 2001, rather belatedly, I published a monograph on the concept of ITSs [10]. Although such ITSs can be regarded as extended structural formulas with colored bonds (par-bonds, out-bonds, and in-bonds), the XƎMTEX system at that time did not support utilities of coloring bonds. It follows that the ITSs contained in this book were drawn manually and pasted on the camera-ready manuscript.

3.2 Drawing by the XyMTEX System

The XyMTEX system was developed and released in 1993 as a LaTeX tool for drawing structural formulas. The manual was published as a book in 1997 [11]. However, it was not until version 4.00 that the XyMTEX system supported the PostScript-compatible mode for drawing structural formulas for high-quality printing [12].

The PostScript-compatible mode was applied to prepare a book for surveying organic compounds for color photography [13]. Along with chemical or mathematical equations, this book contains 480 figures, each of which consists of several structural formulas drawn by the XyMTEX system.

The book published in 2007 deals with a new concept, *mandalas*, which I have proposed as a basis for rationalizing enumeration of three-dimensional structures [14]. This book contains many mathematical equations as well as structural formulas because of its interdisciplinary nature; the mathematical equations were again typeset by the original (LA)TEX utilities, but this time the structural formulas were drawn by the XyMTEX system.

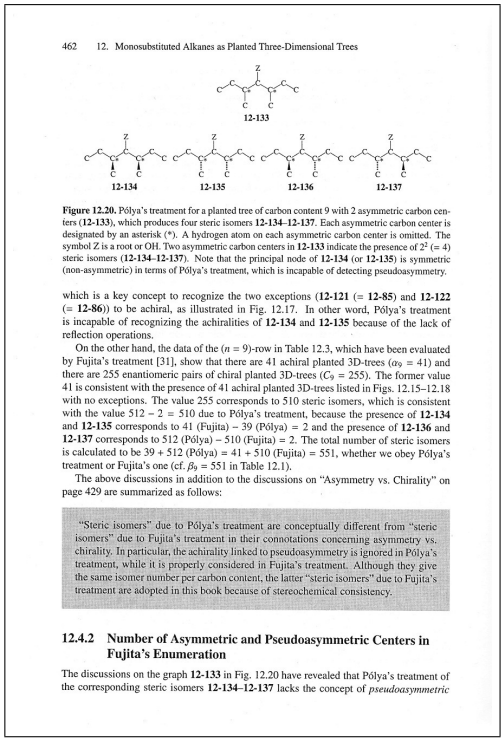
The book published in 2013 is concerned with the *proligand method*, in which I have proposed to enumerate three-dimensional structures [15]. This book indicates that the proligand method for enumerating three-dimensional structures can be degenerated into Pólya's method for enumerating graphs.

A sample page shown in Fig. 1 (page 462 of [15]) contains structural formulas drawn by the XyMTEX system, while another sample page shown in Fig. 2 (page 463 of [15]) contains mathematical equations typeset by the original utilities of the LaTeX system.

These sample pages from [15] demonstrate that the combination of the XyMTEX system with the LaTeX system is an efficient tool for publishing interdisciplinary chemistry/mathematics books.

Moreover, the on-line manual [1] of the XyMTEX system itself provides us with an illustrative example for publishing a book which contains both chemical structural formulas and mathematical equations. For example, several structural formulas drawn by the XyMTEX system are aligned in an `align` environment of the `amsmath` package bundled with the LaTeX system, so as to generate a reaction scheme, as shown in Fig. 3 (page 647 of [1]).

Because XyMTEX version 5.01 supports utilities for coloring structural formulas, the book published in 2001 would be rewritten with maintaining bond colors (par-bonds, out-bonds, and in-bonds). This has been briefly discussed in Section 39.4 of the on-line manual [1].



462 12. Monosubstituted Alkanes as Planted Three-Dimensional Trees

12-133

12-134 12-135 12-136 12-137

Figure 12.20. Pólya's treatment for a planted tree of carbon content 9 with 2 asymmetric carbon centers (**12-133**), which produces four steric isomers **12-134**–**12-137**. Each asymmetric carbon center is designated by an asterisk (*). A hydrogen atom on each asymmetric carbon center is omitted. The symbol Z is a root or OH. Two asymmetric carbon centers in **12-133** indicate the presence of 2^2 (= 4) steric isomers (**12-134**–**12-137**). Note that the principal node of **12-134** (or **12-135**) is symmetric (non-asymmetric) in terms of Pólya's treatment, which is incapable of detecting pseudoasymmetry.

which is a key concept to recognize the two exceptions (**12-121** (= **12-85**) and **12-122** (= **12-86**)) to be achiral, as illustrated in Fig. 12.17. In other word, Pólya's treatment is incapable of recognizing the achiralities of **12-134** and **12-135** because of the lack of reflection operations.

On the other hand, the data of the (n = 9)-row in Table 12.3, which have been evaluated by Fujita's treatment [31], show that there are 41 achiral planted 3D-trees (α_9 = 41) and there are 255 enantiomeric pairs of chiral planted 3D-trees (C_9 = 255). The former value 41 is consistent with the presence of 41 achiral planted 3D-trees listed in Figs. 12.15–12.18 with no exceptions. The value 255 corresponds to 510 steric isomers, which is consistent with the value 512 − 2 = 510 due to Pólya's treatment, because the presence of **12-134** and **12-135** corresponds to 41 (Fujita) − 39 (Pólya) = 2 and the presence of **12-136** and **12-137** corresponds to 512 (Pólya) − 510 (Fujita) = 2. The total number of steric isomers is calculated to be 39 + 512 (Pólya) = 41 + 510 (Fujita) = 551, whether we obey Pólya's treatment or Fujita's one (cf. β_9 = 551 in Table 12.1).

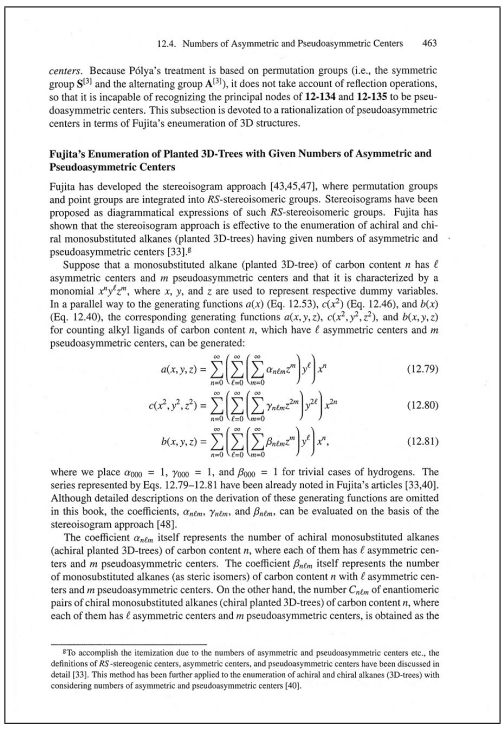
The above discussions in addition to the discussions on "Asymmetry vs. Chirality" on page 429 are summarized as follows:

> "Steric isomers" due to Pólya's treatment are conceptually different from "steric isomers" due to Fujita's treatment in their connotations concerning asymmetry vs. chirality. In particular, the achirality linked to pseudoasymmetry is ignored in Pólya's treatment, while it is properly considered in Fujita's treatment. Although they give the same isomer number per carbon content, the latter "steric isomers" due to Fujita's treatment are adopted in this book because of stereochemical consistency.

12.4.2 Number of Asymmetric and Pseudoasymmetric Centers in Fujita's Enumeration

The discussions on the graph **12-133** in Fig. 12.20 have revealed that Pólya's treatment of the corresponding steric isomers **12-134**–**12-137** lacks the concept of *pseudoasymmetric*

Figure 1: Sample page containing structural formulas drawn by the XyMTEX system (page 462 of [15]).



12.4. Numbers of Asymmetric and Pseudoasymmetric Centers 463

centers. Because Pólya's treatment is based on permutation groups (i.e., the symmetric group $\mathbf{S}^{[3]}$ and the alternating group $\mathbf{A}^{[3]}$), it does not take account of reflection operations, so that it is incapable of recognizing the principal nodes of **12-134** and **12-135** to be pseudoasymmetric centers. This subsection is devoted to a rationalization of pseudoasymmetric centers in terms of Fujita's eneumeration of 3D structures.

Fujita's Enumeration of Planted 3D-Trees with Given Numbers of Asymmetric and Pseudoasymmetric Centers

Fujita has developed the stereoisogram approach [43,45,47], where permutation groups and point groups are integrated into *RS*-stereoisomeric groups. Stereoisograms have been proposed as diagrammatical expressions of such *RS*-stereoisomeric groups. Fujita has shown that the stereoisogram approach is effective to the enumeration of achiral and chiral monosubstituted alkanes (planted 3D-trees) having given numbers of asymmetric and pseudoasymmetric centers [33].[g]

Suppose that a monosubstituted alkane (planted 3D-tree) of carbon content n has ℓ asymmetric centers and m pseudoasymmetric centers and that it is characterized by a monomial $x^n y^\ell z^m$, where x, y, and z are used to represent respective dummy variables. In a parallel way to the generating functions $a(x)$ (Eq. 12.53), $c(x^2)$ (Eq. 12.46), and $b(x)$ (Eq. 12.40), the corresponding generating functions $a(x, y, z)$, $c(x^2, y^2, z^2)$, and $b(x, y, z)$ for counting alkyl ligands of carbon content n, which have ℓ asymmetric centers and m pseudoasymmetric centers, can be generated:

$$a(x,y,z) = \sum_{n=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\alpha_{n\ell m}z^m\right)y^\ell\right)x^n \tag{12.79}$$

$$c(x^2,y^2,z^2) = \sum_{n=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\gamma_{n\ell m}z^{2m}\right)y^{2\ell}\right)x^{2n} \tag{12.80}$$

$$b(x,y,z) = \sum_{n=0}^{\infty}\left(\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty}\beta_{n\ell m}z^m\right)y^\ell\right)x^n, \tag{12.81}$$

where we place α_{000} = 1, γ_{000} = 1, and β_{000} = 1 for trivial cases of hydrogens. The series represented by Eqs. 12.79–12.81 have been already noted in Fujita's articles [33,40]. Although detailed descriptions on the derivation of these generating functions are omitted in this book, the coefficients, $\alpha_{n\ell m}$, $\gamma_{n\ell m}$, and $\beta_{n\ell m}$, can be evaluated on the basis of the stereoisogram approach [48].

The coefficient $\alpha_{n\ell m}$ itself represents the number of achiral monosubstituted alkanes (achiral planted 3D-trees) of carbon content n, where each of them has ℓ asymmetric centers and m pseudoasymmetric centers. The coefficient $\beta_{n\ell m}$ itself represents the number of monosubstituted alkanes (as steric isomers) of carbon content n with ℓ asymmetric centers and m pseudoasymmetric centers. On the other hand, the number $C_{n\ell m}$ of enantiomeric pairs of chiral monosubstituted alkanes (chiral planted 3D-trees) of carbon content n, where each of them has ℓ asymmetric centers and m pseudoasymmetric centers, is obtained as the

[g] To accomplish the itemization due to the numbers of asymmetric and pseudoasymmetric centers etc., the definitions of *RS* -stereogenic centers, asymmetric centers, and pseudoasymmetric centers have been discussed in detail [33]. This method has been further applied to the enumeration of achiral and chiral alkanes (3D-trees) with considering numbers of asymmetric and pseudoasymmetric centers [40].

Figure 2: Sample page containing mathematical equations (page 463 of [15]), which are typeset by the original utilities of the LaTeX system.

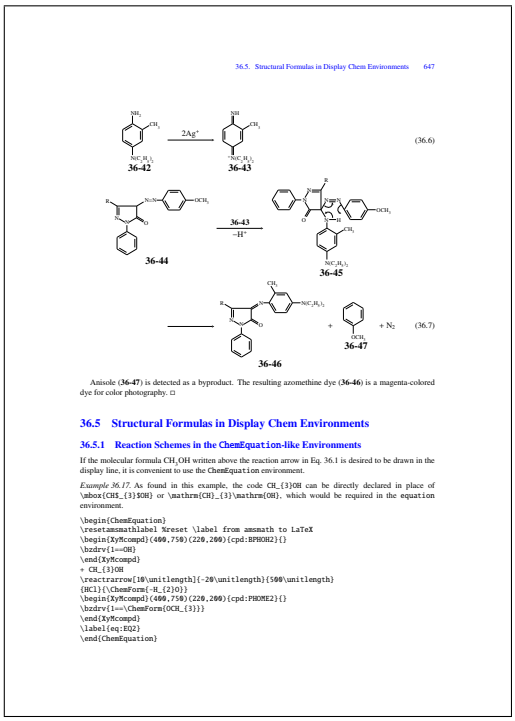


Figure 3: Sample page of the XyMTEX manual [1] (page 647), which contains a reaction scheme drawn by the XyMTEX system and aligned in an `align` environment of the `amsmath` package.

4 Conclusion

As clarified by the publication of the interdisciplinary chemistry/mathematics books described above, the XyMTEX system coupled with the LaTeX system has been proven to be a reliable tool for publishing books of high printing quality which contain structural formulas along with mathematical equations.

References

[1] S. Fujita, "XyMTEX: Reliable Tool for Drawing Chemical Structural Formulas," Shonan Institute of Chemoinformatics and Mathematical Chemistry, Kanagawa (2013), `http://xymtex.com/fujitas3/xymtex/indexe.html`.

[2] S. Fujita, *J. Chem. Inf. Comput. Sci.*, **26**, 205–212 (1986).

[3] S. Fujita, *Theor. Chim. Acta*, **76**, 247–268 (1989).

[4] S. Fujita, *J. Am. Chem. Soc.*, **112**, 3390–3397 (1990).

[5] S. Fujita, *J. Org. Chem.*, **69**, 3158–3165 (2004).

[6] S. Fujita, *Tetrahedron*, **60**, 11629–11638 (2004).

[7] S. Fujita, *Theor. Chem. Acc.*, **113**, 73–79 (2005).

[8] S. Fujita, *J. Math. Chem.*, **42**, 481–534 (2007).

[9] S. Fujita, "Symmetry and Combinatorial Enumeration in Chemistry," Springer-Verlag, Berlin-Heidelberg (1991).

[10] S. Fujita, "Computer-Oriented Representation of Organic Reactions," Yoshioka-Shoten, Kyoto (2001).

[11] S. Fujita, "XyMTEX — Typesetting Chemical Structural Formulas," Addison-Wesley Japan, Tokyo (1997).

[12] S. Fujita, *J. Comput. Chem. Jpn.*, **4**, 69–78 (2005).

[13] S. Fujita, "Organic Chemistry of Photography," Springer-Verlag, Berlin-Heidelberg (2004).

[14] S. Fujita, "Diagrammatical Approach to Molecular Symmetry and Enumeration of Stereoisomers," University of Kragujevac, Faculty of Science, Kragujevac (2007).

[15] S. Fujita, "Combinatorial Enumeration of Graphs, Three-Dimensional Structures, and Chemical Compounds," University of Kragujevac, Faculty of Science, Kragujevac (2013).

◇ Shinsaku Fujita
Shonan Institute of Chemoinformatics and Mathematical Chemistry
`http://xymtex.com`