

Development of

# TeXShop

- The Past and the Future -

Yusuke Terada

Tetsuryokukai (鉄緑会)

# SUMMARY

---

1. The history of TeXShop
2. TeXShop's features equipped for editing Japanese documents
3. The future of TeXShop

# WHAT IS TEXSHOP?

---

- TeX editor and previewer for Mac OS X
- Open source software (GPL)
- Developed by Richard Koch

# TEXSHOP

```
133
134 \title{The Spectral Structure of a Simple Model for Quantum Wire}
135 \author{Yusuke TERADA\thanks{Graduate School of Mathematical Science, University of Tokyo}}
136 \date{}
137
138 \begin{document}
139 \maketitle
140 \abstract{}
141 We consider a simple model for the system which is called "quantum wire" in solid state
142 physics.
143 Our model is one-particle quantum-mechanical system on  $\Omega \subset \mathbb{R}^2$ ,
144 which is the union of the half-plane  $\Omega_+$  with a constant magnetic field and the
145 half-infinite strap-shaped domain
146  $\Omega_-$  with no scalar or vector potential.
147 We study the spectral structure of the Hamiltonian in our model by the
148 scattering theory and Mourre method.
149 }
150 \section{Introduction}
151 We consider the quantum-mechanical motion of a positively charged particle moving
152 in a domain  $\Omega \subset \mathbb{R}^2$  with a magnetic field.
153 If  $\Omega = \mathbb{R}^2$  and the magnetic field is constant, the result is well-known.
154 We review this case. We denote the Hamiltonian by
155 \begin{align*}
156 H_0 &= (-i\nabla - A)^2 \\
157 &= -\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} + \left(-i\frac{\partial}{\partial x_2} - Bx_1\right)^2,
158 \end{align*}
159 where  $B > 0$  is the constant magnetic field and  $A(x_1, x_2) = (0, Bx_1)$  is the vector
160 potential (Landau gauge).
161 By the Fourier transformation with respect to  $x_2$ , we have
162 \begin{align*}
163 H_0 &= \int_{\mathbb{R}} H(k) dk \quad (\text{direct integral decomposition}), \\
164 H(k) &= -\frac{d^2}{dx_1^2} + (k - Bx_1)^2.
165 \end{align*}
166 For each  $k$ ,  $H(k)$  is the Hamiltonian of the one-dimensional harmonic oscillator, though
167 its center is not the origin in general. Hence the spectrum of  $H(k)$  is
168 \begin{align*}
169 \sigma(H(k)) = \sigma_{\text{pp}}(H(k)) = \{(2n-1)B \mid n=1, 2, \dots\} = (2\mathbb{N}-1)B.
170 \end{align*}
171 Since this holds for each  $k$ , we have
172 \begin{align*}
173 \sigma(H_0) = \sigma_{\text{pp}}(H_0) = (2\mathbb{N}-1)B \quad (\text{Landau levels}).
174 \end{align*}
175 This means that all states are bound states.
176 \vspace{5mm}
```

Source Window

$\Omega \subset \mathbb{R}^2$ , which is the union of the half-plane  $\Omega_+$  with a constant magnetic field and the half-infinite strap-shaped domain  $\Omega_-$  with no scalar or vector potential. We study the spectral structure of the Hamiltonian in our model by the scattering theory and Mourre method.

## 1 Introduction

We consider the quantum-mechanical motion of a positively charged particle moving in a domain  $\Omega \subset \mathbb{R}^2$  with a magnetic field. If  $\Omega = \mathbb{R}^2$  and the magnetic field is constant, the result is well-known. We review this case. We denote the Hamiltonian by

$$H_0 = (-i\nabla - A)^2 = -\frac{\partial^2}{\partial x_1^2} + \left(-i\frac{\partial}{\partial x_2} - Bx_1\right)^2,$$

where  $B > 0$  is the constant magnetic field and  $A(x_1, x_2) = (0, Bx_1)$  is the vector potential (Landau gauge).

By the Fourier transformation with respect to  $x_2$ , we have

$$H_0 = \int_{\mathbb{R}} H(k) dk \quad (\text{direct integral decomposition}),$$
$$H(k) = -\frac{d^2}{dx_1^2} + (k - Bx_1)^2.$$

For each  $k$ ,  $H(k)$  is the Hamiltonian of the one-dimensional harmonic oscillator, though its center is not the origin in general. Hence the spectrum of  $H(k)$  is

$$\sigma(H(k)) = \sigma_{\text{pp}}(H(k)) = \{(2n-1)B \mid n=1, 2, \dots\} = (2\mathbb{N}-1)B.$$

Since this holds for each  $k$ , we have  $\sigma(H_0) = \sigma_{\text{pp}}(H_0) = (2\mathbb{N}-1)B$  (Landau levels). This means that all states are bound states.

However, if we restrict  $\Omega$  to the half-plane  $(0, \infty) \times \mathbb{R}$  and impose the Dirichlet boundary condition, the result changes greatly.

---

<sup>†</sup>Graduate School of Mathematical Science, University of Tokyo

Preview Window

Command Completion      Key Bindings      Spell Checking  
Templates      Parenthesis Matching      SyncTeX  
AppleScript Support      Macros      Matrix Input Panel  
Resume      Symbol Input Panel      Versions  
Syntax Coloring      Automatic Saving      Localizations  
\UTF / \CID export      Unicode Normalization  
Multi-Display Support      Regular Expression  
Split Window      Show Invisible Characters  
Place Holder      BibTeX  
Full Screen      Switching TeX Engines      Page Layout

# HISTORY

---

- 1985  
Steve Jobs got fired from Apple and founded NeXT.
- 1989 Sep.  
NeXTSTEP 1.0 was released. TeXview 1.0 was released by Tom Rokicki.
- 1996  
Apple bought NeXT and Steve Jobs returned to Apple.
- 1999  
Koch started developing TeXShop for Mac OS X beta, inspired by TeXview.

# HISTORY

---

YEAR	Mac OS X	Feature	TeXShop	
2001	10	Cheetah	1	
2001	10.1	Puma		
2002	10.2	Jaguar		
2003	10.3	Panther		
2005	10.4	Tiger	PDF Kit	2
2007	10.5	Leopard		
2009	10.6	Snow		
2011	10.7	Lion	Resume	3
2012	10.8	Mountain		
2013	10.9	Mavericks		

---

---


**TEXSHOP  
AND  
JAPANESE DEVELOPERS**



# BACKSLASH - YEN MARK PROBLEM

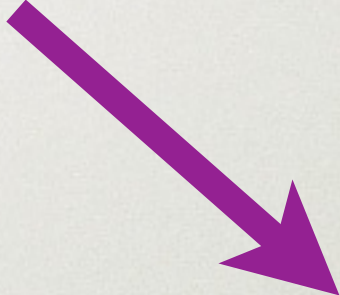
---

Press Backslash Key  
on Japanese Keyboard



¥

U+00A5  
(default)



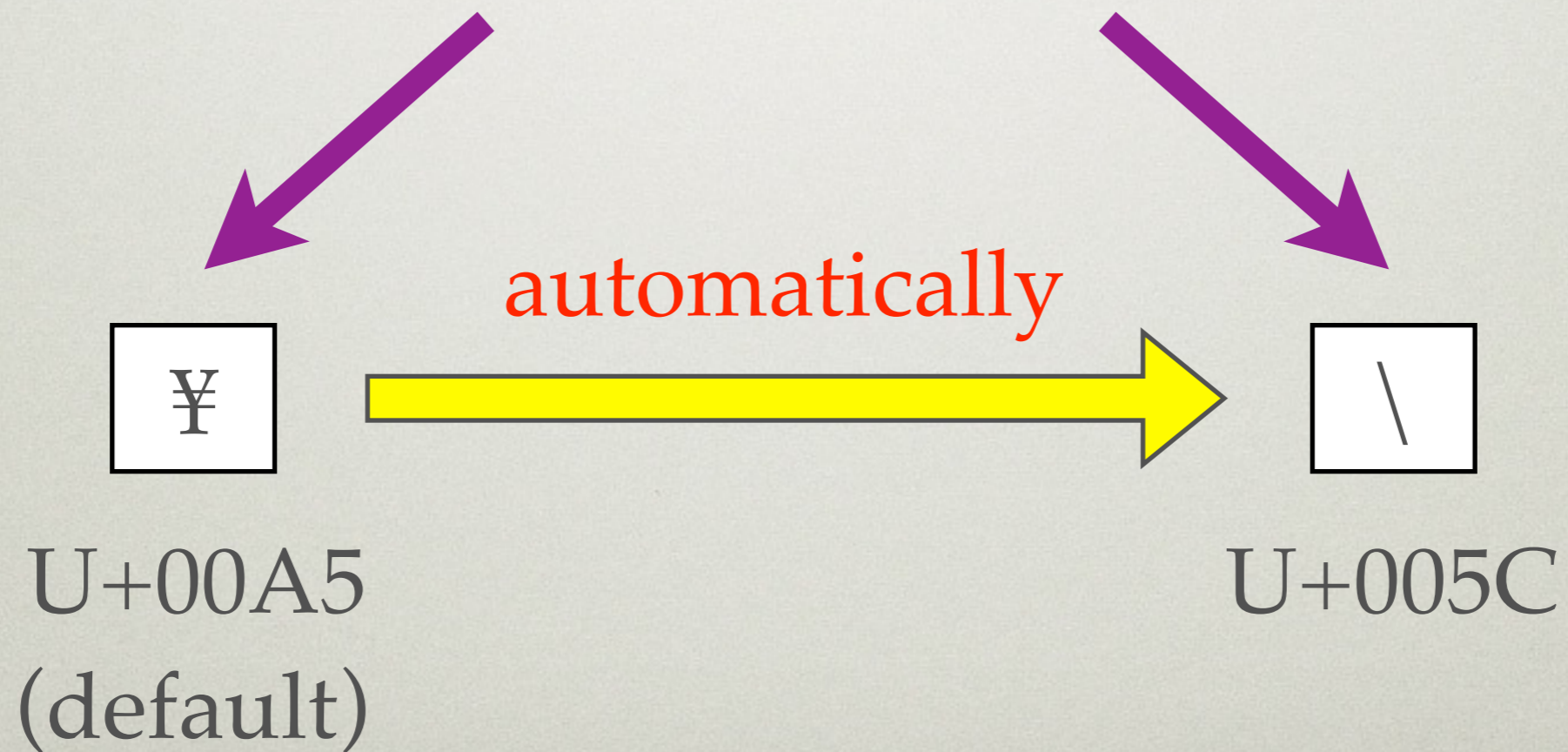
\

U+005C

# BACKSLASH - YEN MARK PROBLEM

---

Press Backslash Key  
on Japanese Keyboard



# JAPANESE COLLABORATORS

---

- Mitsuhiro Shishikura
- Seiji Zenitani
- Makoto Inoue
- Yu Itoh
- Koichi Inoue
- Yoshihisa Okazaki
- Tomoaki Okayama
- Isao Sonobe

---

**TEXSHOP'S FEATURES  
EQUIPPED FOR EDITING  
JAPANESE DOCUMENTS**

# FULL-WIDTH(ZENKAKU) SPACE

---

```
\begin{enumerate}  
\item first  
\item second  
\end{enumerate}
```

# FULL-WIDTH(ZENKAKU) SPACE

---

```
\begin{enumerate}
```

```
\item first
```

```
\item second
```

```
\end{enumerate}
```

➔ Successfully compiled

➔ ERROR: Undefined

control sequence

# FULL-WIDTH (ZENKAKU) SPACE

---

Ordinary Space

```
\begin{enumerate}
\item first
\item second
\end{enumerate}
```

→ Successfully compiled

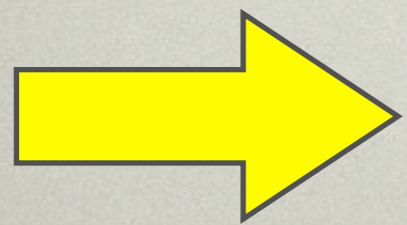
→ ERROR: Undefined  
control sequence

Zenkaku Space

# FULL-WIDTH(ZENKAKU) SPACE

---

```
\begin{enumerate}  
\item_ first  
\item□ second  
\end{enumerate}
```



Easy to distinguish



# JAPANESE ENCODINGS

---

- Shift-JIS
- JIS (ISO-2022-JP)
- EUC-JP
- UTF-8
- UTF-16            etc...

# EMOJI

(Digression)



# EMOJI

(Digression)

```
taylor — bash — 33x10  
$ echo 😄 > 😭  
$ cat 😭  
😄  
$ ls 😭*  
😭 😭👍 😭🌟❌  
$ █
```

# CIRCLED NUMBERS

---

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

⑪ ⑫ ⑬ ⑭ ⑮ ⑯ ⑰ ⑱ ⑲ ⑳

㉑ ㉒ ㉓ ㉔ ㉕ ㉖ ㉗ ㉘ ㉙ ㉚

㉛ ㉜ ㉝ ㉞ ㉟ ㊱ ㊲ ㊳ ㊴ ㊵

㊶ ㊷ ㊸ ㊹ ㊺ ㊻ ㊼ ㊽ ㊾ ㊿

# CIRCLED NUMBERS

---

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

⑪ ⑫ ⑬ ⑭ ⑮ ⑯ ⑰ ⑱ ⑲ ⑳

Lost

Shift-JIS / EUC-JP

# UTF/OTF PACKAGE

---

①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳

\ajMaru{21}\ajMaru{22}\ajMaru{23}\ajMaru{24}\ajMaru{25}

\ajMaru{26}\ajMaru{27}\ajMaru{28}\ajMaru{29}\ajMaru{30}

\ajMaru{31}\ajMaru{32}\ajMaru{33}\ajMaru{34}\ajMaru{35}

\ajMaru{36}\ajMaru{37}\ajMaru{38}\ajMaru{39}\ajMaru{40}

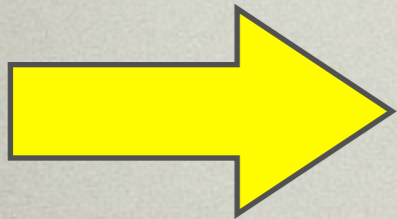
\ajMaru{41}\ajMaru{42}\ajMaru{43}\ajMaru{44}\ajMaru{45}

\ajMaru{46}\ajMaru{47}\ajMaru{48}\ajMaru{49}\ajMaru{50}

# UTF/OTF PACKAGE

①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳  
\ajMaru{21}\ajMaru{22}\ajMaru{23}\ajMaru{24}\ajMaru{25}  
\ajMaru{26}\ajMaru{27}\ajMaru{28}\ajMaru{29}\ajMaru{30}  
\ajMaru{31}\ajMaru{32}\ajMaru{33}\ajMaru{34}\ajMaru{35}  
\ajMaru{36}\ajMaru{37}\ajMaru{38}\ajMaru{39}\ajMaru{40}  
\ajMaru{41}\ajMaru{42}\ajMaru{43}\ajMaru{44}\ajMaru{45}  
\ajMaru{46}\ajMaru{47}\ajMaru{48}\ajMaru{49}\ajMaru{50}

Compiled



①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰  
⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟  
㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿

# UTF/OTF PACKAGE

---

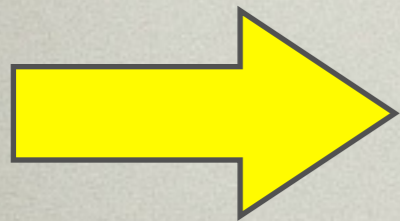
①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑳  
\UTF{3251}\UTF{3252}\UTF{3253}\UTF{3254}\UTF{3255}  
\UTF{3256}\UTF{3257}\UTF{3258}\UTF{3259}\UTF{325A}  
\UTF{325B}\UTF{325C}\UTF{325D}\UTF{325E}\UTF{325F}  
\UTF{32B1}\UTF{32B2}\UTF{32B3}\UTF{32B4}\UTF{32B5}  
\UTF{32B6}\UTF{32B7}\UTF{32B8}\UTF{32B9}\UTF{32BA}  
\UTF{32BB}\UTF{32BC}\UTF{32BD}\UTF{32BE}\UTF{32BF}



# UTF/OTF PACKAGE

①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳  
\UTF{3251}\UTF{3252}\UTF{3253}\UTF{3254}\UTF{3255}  
\UTF{3256}\UTF{3257}\UTF{3258}\UTF{3259}\UTF{325A}  
\UTF{325B}\UTF{325C}\UTF{325D}\UTF{325E}\UTF{325F}  
\UTF{32B1}\UTF{32B2}\UTF{32B3}\UTF{32B4}\UTF{32B5}  
\UTF{32B6}\UTF{32B7}\UTF{32B8}\UTF{32B9}\UTF{32BA}  
\UTF{32BB}\UTF{32BC}\UTF{32BD}\UTF{32BE}\UTF{32BF}

Compiled



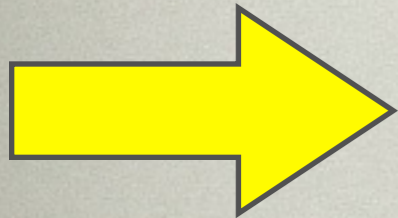
①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰  
⑱⑲⑳㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟  
㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿

# UTF-EXPORT

①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳  
㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺  
㊻㊼㊽㊾㊿

*In Memory*

*Saved*



```
\UTF{2460}\UTF{2461}\UTF{2462}\UTF{2463}\UTF{2464}\UTF{2465}  
\UTF{2466}\UTF{2467}\UTF{2468}\UTF{2469}\UTF{246A}\UTF{246B}  
\UTF{246C}\UTF{246D}\UTF{246E}\UTF{246F}\UTF{2470}\UTF{2471}  
\UTF{2472}\UTF{2473}\UTF{3251}\UTF{3252}\UTF{3253}\UTF{3254}  
\UTF{3255}\UTF{3256}\UTF{3257}\UTF{3258}\UTF{3259}\UTF{325A}  
\UTF{325B}\UTF{325C}\UTF{325D}\UTF{325E}\UTF{325F}\UTF{32B1}  
\UTF{32B2}\UTF{32B3}\UTF{32B4}\UTF{32B5}\UTF{32B6}\UTF{32B7}  
\UTF{32B8}\UTF{32B9}\UTF{32BA}\UTF{32BB}\UTF{32BC}\UTF{32BD}  
\UTF{32BE}\UTF{32BF}
```

# UTF-EXPORT

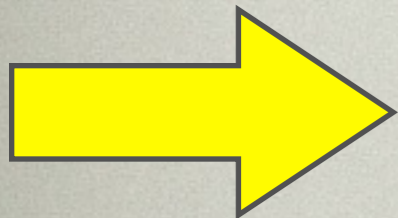
---

```
\UTF{2460}\UTF{2461}\UTF{2462}\UTF{2463}\UTF{2464}\UTF{2465}  
\UTF{2466}\UTF{2467}\UTF{2468}\UTF{2469}\UTF{246A}\UTF{246B}  
\UTF{246C}\UTF{246D}\UTF{246E}\UTF{246F}\UTF{2470}\UTF{2471}  
\UTF{2472}\UTF{2473}\UTF{3251}\UTF{3252}\UTF{3253}\UTF{3254}  
\UTF{3255}\UTF{3256}\UTF{3257}\UTF{3258}\UTF{3259}\UTF{325A}  
\UTF{325B}\UTF{325C}\UTF{325D}\UTF{325E}\UTF{325F}\UTF{32B1}  
\UTF{32B2}\UTF{32B3}\UTF{32B4}\UTF{32B5}\UTF{32B6}\UTF{32B7}  
\UTF{32B8}\UTF{32B9}\UTF{32BA}\UTF{32BB}\UTF{32BC}\UTF{32BD}  
\UTF{32BE}\UTF{32BF}
```

# UTF-EXPORT

```
\UTF{2460}\UTF{2461}\UTF{2462}\UTF{2463}\UTF{2464}\UTF{2465}  
\UTF{2466}\UTF{2467}\UTF{2468}\UTF{2469}\UTF{246A}\UTF{246B}  
\UTF{246C}\UTF{246D}\UTF{246E}\UTF{246F}\UTF{2470}\UTF{2471}  
\UTF{2472}\UTF{2473}\UTF{3251}\UTF{3252}\UTF{3253}\UTF{3254}  
\UTF{3255}\UTF{3256}\UTF{3257}\UTF{3258}\UTF{3259}\UTF{325A}  
\UTF{325B}\UTF{325C}\UTF{325D}\UTF{325E}\UTF{325F}\UTF{32B1}  
\UTF{32B2}\UTF{32B3}\UTF{32B4}\UTF{32B5}\UTF{32B6}\UTF{32B7}  
\UTF{32B8}\UTF{32B9}\UTF{32BA}\UTF{32BB}\UTF{32BC}\UTF{32BD}  
\UTF{32BE}\UTF{32BF}
```

Loaded



①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯⑰⑱⑲⑳  
㉑㉒㉓㉔㉕㉖㉗㉘㉙㉚㉛㉜㉝㉞㉟㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺  
㊻㊼㊽㊾㊿

In Memory

# DAKUTEN / HAN-DAKUTEN

---

は

ha



ば

ba

ぱ

pa

# DAKUTEN / HAN-DAKUTEN

---

は

ha



は

ba


Dakuten

は

pa

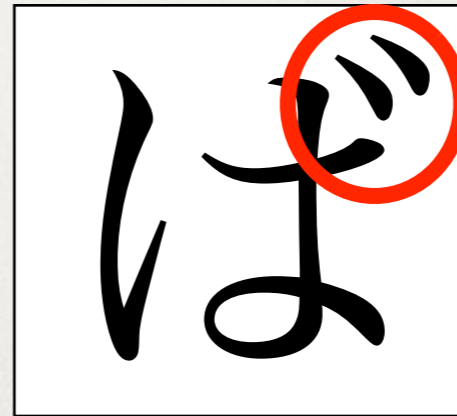
# DAKUTEN / HAN-DAKUTEN

---



は

ha



は

ba

Dakuten



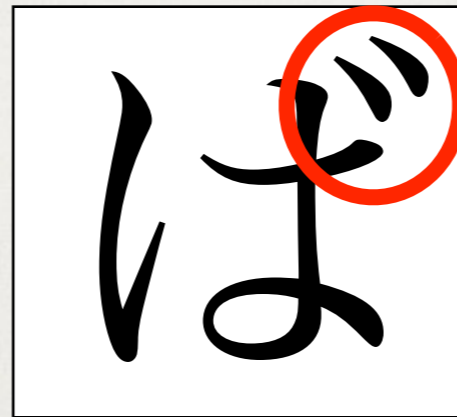
は

pa

Han-Dakuten

# DAKUTEN

---

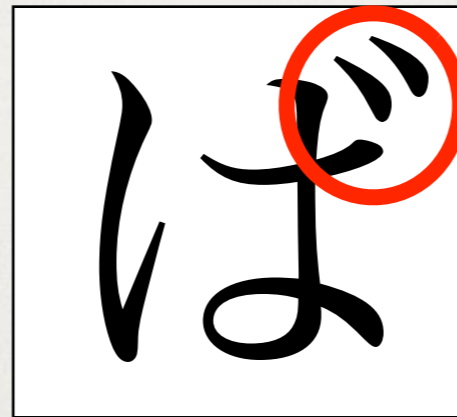


Dakuten



# DAKUTEN

---



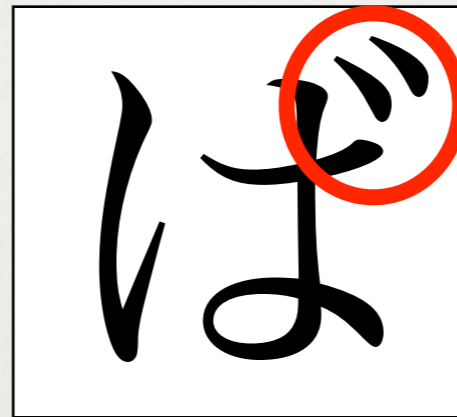
Dakuten



(U+3070) Hiragana letter “ba”


# DAKUTEN

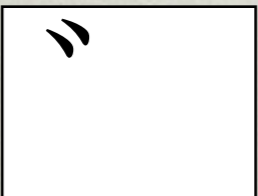
---



Dakuten

\*  (U+3070) Hiragana letter “ba”

\*  (U+306F) Hiragana letter “ha”

+  (U+3099) Combining katakana-hiragana  
voiced sound mark

# DAKUTEN ISSUE

---

```
\def\ほげ{hoge}  
\ほげ  
\ほげ
```

# DAKUTEN ISSUE

---

`\def\ほげ{hoge}`

`\ほげ` → Successfully compiled

`\ほげ` → ERROR: Undefined control sequence

# DAKUTEN ISSUE

---

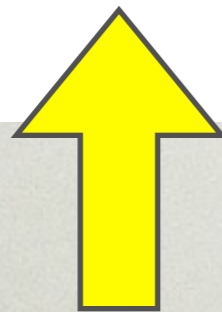
\def\ほげ{hoge}

\ほげ

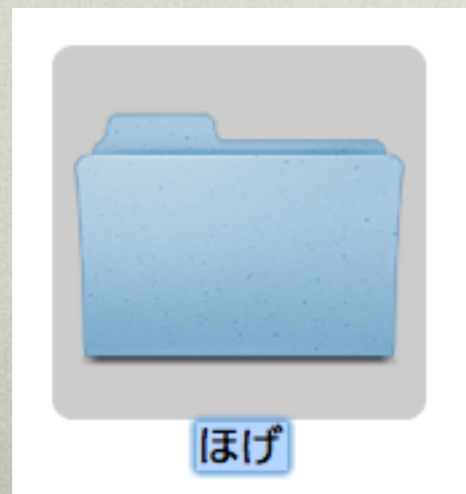
→ Successfully compiled

\ほげ

→ ERROR: Undefined control sequence



Copy & Paste



Filename on Finder

# UNICODE NORMALIZATION

---

- NFC : Normalization Form C  
(Composition)
- NFD : Normalization Form D  
(Decomposition)

# NFD

---

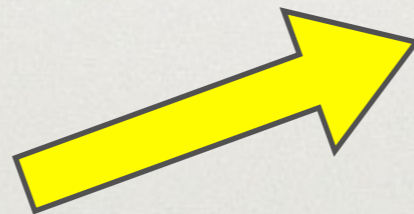
ほ	げ
---	---

decomposed



ほ	け	゛
---	---	---

ほ	け	゛
---	---	---



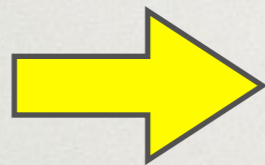
decomposed  
(unchanged)

# NFC

---

decomposed

ほ	げ
---	---



ほ	げ	〃
---	---	---

composed



ほ	げ	〃
---	---	---



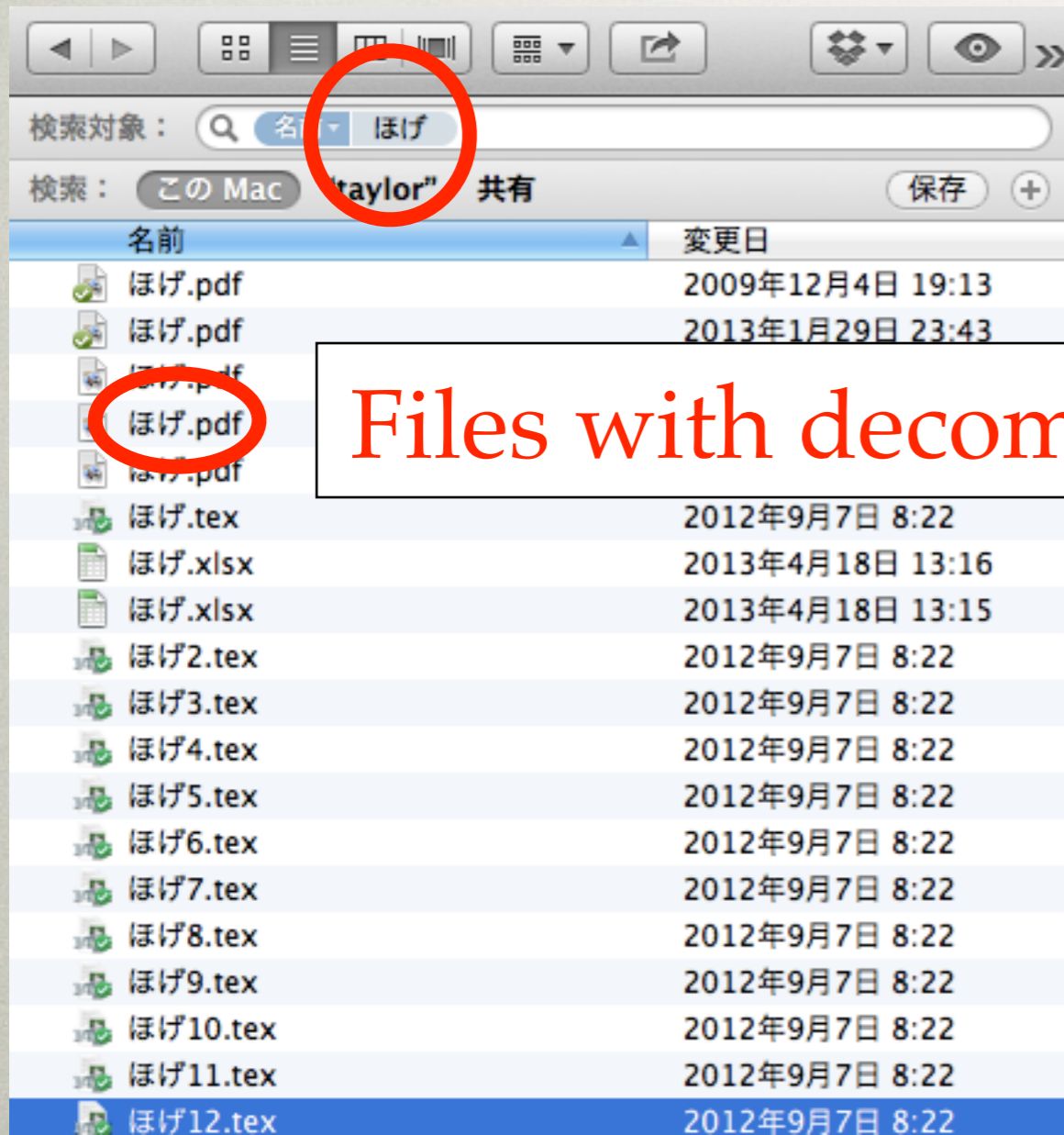
composed

ほ	げ
---	---



# UNICODE NORMALIZATION

Search a filename with composed-form characters



Files with decomposed-form names hit.

---

**LOST IN  
NORMALIZATION**

**MY NAME**

---

# MY NAME

---

- English: Yusuke TERADA
- Hiragana: てらだ ゆうすけ
- Kanji: 寺田 侑祐

# MY NAME

---

- English: Yusuke TERADA
- Hiragana: てらだ ゆうすけ
- Kanji: 寺田 侑祐

---

**DEMONSTRATION**  
**(SAFARI / MAIL)**

# CJK COMPATIBILITY

## IDEOGRAPHS

欄 → 欄	廊 → 廊	朗 → 朗	虜 → 虜	殺 → 殺	類 → 類
隆 → 隆	塚 → 塚	熙 → 熙	猪 → 猪	神 → 神	祥 → 祥
福 → 福	蘊 → 蘊	諸 → 諸	都 → 都	侮 → 侮	僧 → 僧
免 → 免	勉 → 勉	勤 → 勤	卑 → 卑	喝 → 喝	嘆 → 嘆
器 → 器	塤 → 塤	墨 → 墨	層 → 層	屮 → 屮	悔 → 悔
慨 → 慨	憎 → 憎	懲 → 懲	敏 → 敏	既 → 既	暑 → 暑
梅 → 梅	海 → 海	渚 → 渚	漢 → 漢	煮 → 煮	𠂇 → 𠂇
琢 → 琢	碑 → 碑	社 → 社	祉 → 祉	祈 → 祈	祐 → 祐
祖 → 祖	祝 → 祝	禍 → 禍	禎 → 禎	穀 → 穀	突 → 突
節 → 節	練 → 練	縉 → 縉	繁 → 繁	署 → 署	者 → 者
臭 → 臭	𠂇 → 𠂇	𠂇 → 𠂇	著 → 著	褐 → 褐	視 → 視
謁 → 謁	謹 → 謹	賓 → 賓	贈 → 贈	讠 → 讠	逸 → 逸
難 → 難	響 → 響	頻 → 頻			

# CJK COMPATIBILITY

## IDEOGRAPHS

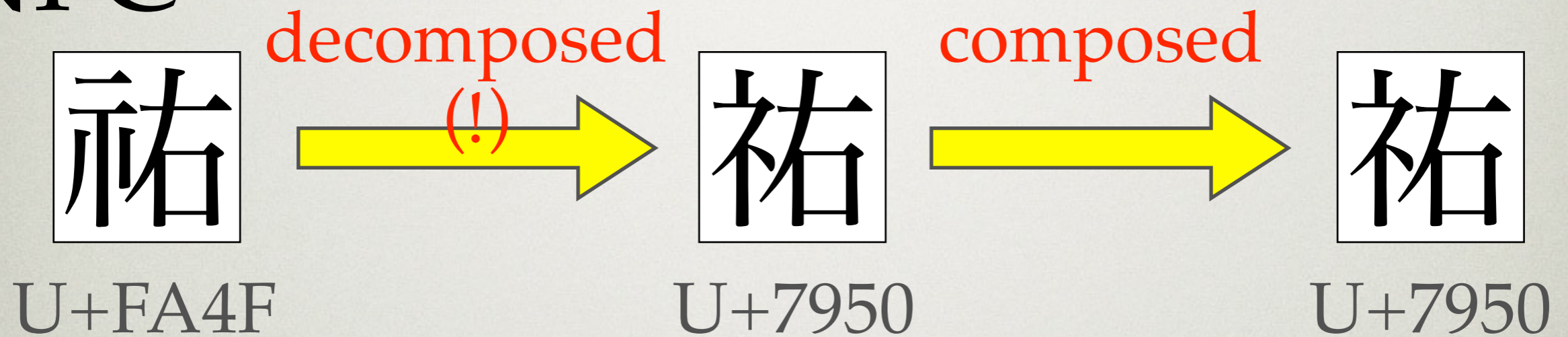
欄 → 欄	廊 → 廊	朗 → 朗	虜 → 虜	殺 → 殺	類 → 類
隆 → 隆	塚 → 塚	熙 → 熙	猪 → 猪	神 → 神	祥 → 祥
福 → 福	蘊 → 蘊	諸 → 諸	都 → 都	侮 → 侮	僧 → 僧
免 → 免	勉 → 勉	勤 → 勤	卑 → 卑	喝 → 喝	嘆 → 嘆
器 → 器	塤 → 塤	墨 → 墨	層 → 層	屮 → 屮	悔 → 悔
慨 → 慨	憎 → 憎	懲 → 懲	敏 → 敏	既 → 既	暑 → 暑
梅 → 梅	海 → 海	渚 → 渚	漢 → 漢	煮 → 煮	𠂇 → 𠂇
琢 → 琢	碑 → 碑	社 → 社	祉 → 祉	祈 → 祈	祐 → 祐
祖 → 祖	祝 → 祝	禍 → 禍	禎 → 禎	穀 → 穀	突 → 突
節 → 節	練 → 練	縉 → 縉	繁 → 繁	署 → 署	者 → 者
臭 → 臭	𠂇 → 𠂇	𠂇 → 𠂇	著 → 著	褐 → 褐	視 → 視
謁 → 謁	謹 → 謹	賓 → 賓	贈 → 贈	讎 → 讎	逸 → 逸
難 → 難	響 → 響	頻 → 頻			



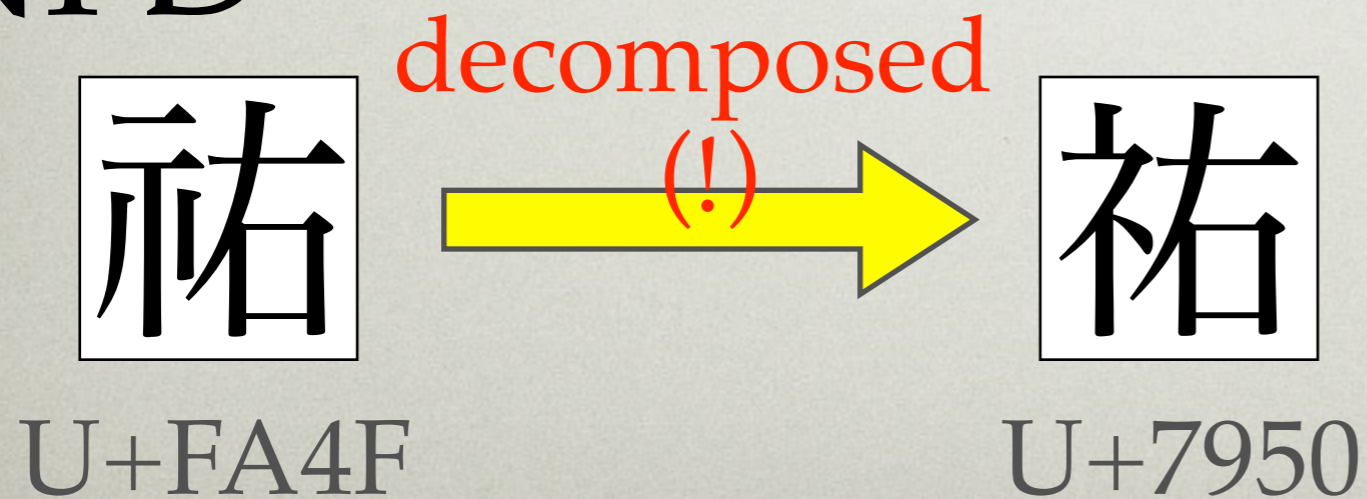
# LOST IN NORMALIZATION

---

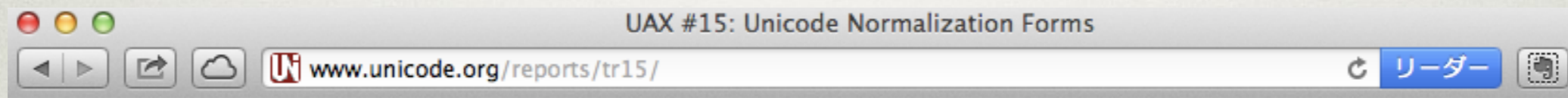
NFC



NFD



# COMPOSITION EXCLUSION



## 5 Composition Exclusion Table

There are four classes of canonically decomposable characters that are excluded from composition:

- Script-specifics:** canonically decomposable characters that are generally not the preferred form for particular scripts
  - These *cannot* be computed from information in the Unicode Character Database.
  - An example is U+0958 (क़) DEVANAGARI LETTER QA.
- Post composition version:** canonically decomposable characters that are added after Unicode 3.0 [Unicode3.0] and whose decompositions exist in prior versions of Unicode. This set will be updated with each subsequent version of Unicode. For more information, see Section 3, [Versioning and Stability](#).
  - These *cannot* be computed from information in the Unicode Character Database.
  - An example is U+2ADC (𐤛) FORKING.
- Singletons:** canonically decomposable characters having decompositions that consist of a single character.
  - These *are* computed from information in the Unicode Character Database.
  - An example is U+2126 (Ω) OHM SIGN.
- Non-starter decompositions:** characters with expanding canonical decompositions (that is, those which canonically decompose to a sequence of characters instead of a single character), such that (A) the character not a Starter, or (B) the character's canonical decomposition starts with a character that is not a Starter. (Note that a "Starter" is any character with a zero combining class).
  - These *are* computed from information in the Unicode Character Database.
  - An example is U+0344 ( ¨ ) COMBINING GREEK DIALYTIKA TONOS.

Two characters may have the same canonical decomposition in the Unicode Character Database.

Table 5. Same Canonical Decomposition

Source	Same Decomposition
212B (Å) ANGSTROM SIGN	0041 (A) LATIN CAPITAL LETTER A + 030A (◌̆) COMBINING RING ABOVE
00C5 (Å) LATIN CAPITAL LETTER A WITH RING ABOVE	

Website of Unicode Consortium

# COMPOSITION EXCLUSION

```
- (NSString*)normalizedStringWithModifiedNFC
{
    NSString *pattern = @"([\x{0340}\x{0341}\x{0343}\x{0344}\x{0374}\x{037E}\x{0387}\x{0958}-\x{095F}\x{09DC}\x{09DD}\x{09DF}\x{0A33}\x{0A36}\x{0A59}-\x{0A5B}\x{0A5E}\x{0B5C}\x{0B5D}\x{0F43}\x{0F4D}\x{0F52}\x{0F57}\x{0F5C}\x{0F69}\x{0F73}\x{0F75}\x{0F76}\x{0F78}\x{0F81}\x{0F93}\x{0F9D}\x{0FA2}\x{0FA7}\x{0FAC}\x{0FB9}\x{1F71}\x{1F73}\x{1F75}\x{1F77}\x{1F79}\x{1F7B}\x{1F7D}\x{1F8B}\x{1FBE}\x{1FC9}\x{1FCB}\x{1FD3}\x{1FDB}\x{1FE3}\x{1FEB}\x{1FEE}\x{1FEF}\x{1FF9}\x{1FFB}\x{1FFD}\x{2000}\x{2001}\x{2126}\x{212A}\x{212B}\x{2329}\x{232A}\x{2ADC}\x{F900}-\x{FA0D}\x{FA10}\x{FA12}\x{FA15}-\x{FA1E}\x{FA20}\x{FA22}\x{FA25}\x{FA26}\x{FA2A}-\x{FA6D}\x{FA70}-\x{FAD9}\x{FB1D}\x{FB1F}\x{FB2A}-\x{FB36}\x{FB38}-\x{FB3C}\x{FB3E}\x{FB40}\x{FB41}\x{FB43}\x{FB44}\x{FB46}-\x{FB4E}\x{1D15E}-\x{1D164}\x{1D1BB}-\x{1D1C0}\x{2F800}-\x{2FA1D}])*([\x{0340}\x{0341}\x{0343}\x{0344}\x{0374}\x{037E}\x{0387}\x{0958}-\x{095F}\x{09DC}\x{09DD}\x{09DF}\x{0A33}\x{0A36}\x{0A59}-\x{0A5B}\x{0A5E}\x{0B5C}\x{0B5D}\x{0F43}\x{0F4D}\x{0F52}\x{0F57}\x{0F5C}\x{0F69}\x{0F73}\x{0F75}\x{0F76}\x{0F78}\x{0F81}\x{0F93}\x{0F9D}\x{0FA2}\x{0FA7}\x{0FAC}\x{0FB9}\x{1F71}\x{1F73}\x{1F75}\x{1F77}\x{1F79}\x{1F7B}\x{1F7D}\x{1F8B}\x{1FBE}\x{1FC9}\x{1FCB}\x{1FD3}\x{1FDB}\x{1FE3}\x{1FEB}\x{1FEE}\x{1FEF}\x{1FF9}\x{1FFB}\x{1FFD}\x{2000}\x{2001}\x{2126}\x{212A}\x{212B}\x{2329}\x{232A}\x{2ADC}\x{F900}-\x{FA0D}\x{FA10}\x{FA12}\x{FA15}-\x{FA1E}\x{FA20}\x{FA22}\x{FA25}\x{FA26}\x{FA2A}-\x{FA6D}\x{FA70}-\x{FAD9}\x{FB1D}\x{FB1F}\x{FB2A}-\x{FB36}\x{FB38}-\x{FB3C}\x{FB3E}\x{FB40}\x{FB41}\x{FB43}\x{FB44}\x{FB46}-\x{FB4E}\x{1D15E}-\x{1D164}\x{1D1BB}-\x{1D1C0}\x{2F800}-\x{2FA1D}])*";
    NSRegularExpression *regexp = [NSRegularExpression regularExpressionWithPattern:pattern options:0 error:nil];
    NSMutableString *result = [NSMutableString stringWithCapacity:0];
    [regexp enumerateMatchesInString:self options:0 range:NSMakeRange(0, [self length]) usingBlock:^(NSTextCheckingResult *match, NSMatchingFlags flags, BOOL *stop){
        [result appendFormat:@"%@@@%",
         [self substringWithRange:[match rangeAtIndex:1]],
         [[self substringWithRange:[match rangeAtIndex:2]] precomposedStringWithCanonicalMapping],
         [self substringWithRange:[match rangeAtIndex:3]]
        ];
    }];
    return result;
}
```

# COMPOSITION EXCLUSION

---

```
\def\ほげ{hoge}
```

```
\ほげ % input directly
```

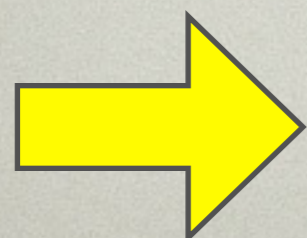
```
\ほげ % paste from Finder
```

```
侑祐 % U+7950
```

```
侑祐 % U+FA4F
```

Normalize

Do not  
normalize



Compiled as users expected

---

**THE FUTURE OF  
TEXSHOP**

# THE FUTURE OF TEXSHOP

---

*“The design goal for TeXShop is simple: lots of space for the user’s work, almost no space for interface and buttons. TeXShop should be in the background ready to help, but never in the way. What is important is the mathematics, physics, philosophy, economics, or history typed by the user. The highest praise is when the user didn’t much notice our program.”*