

pT_EX マニュアル

日本語 T_EX 開発コミュニティ*

version p3.7.2, 2017 年 10 月 12 日

目次

1	禁則	2
2	文字間のスペース	3
3	組方向	5
4	ベースライン補正	7
5	和文フォント	7
6	文字コード変換, 漢数字	8
7	長さ単位	9

* <https://texjp.org>, e-mail: [issue\(at\)texjp.org](mailto:issue(at)texjp.org)

1 禁則

欧文と和文の処理で見かけ上最も大きな違いは、行分割処理であろう。

- 欧文中での行分割は、ハイフネーション処理等の特別な場合を除いて、単語中すなわち連続する文字列中はブレイクポイントとして選択されない。
- 和文中では禁則（行頭禁則と行末禁則）の例外を除いて、全ての文字間がブレイクポイントになり得る。

しかし、pTeX の行分割は TeX 内部の処理からさほど大きな変更を加えられてはいない。というのも、TeX のペナルティ^{*1}という概念を和文の禁則処理にも適用しているためである。

行頭禁則と行末禁則を実現する方法として、pTeX では「禁則テーブル」が用意されている。このテーブルには

- 禁則文字
- その文字に対応するペナルティ値
- ペナルティの挿入位置（その文字の前に挿入するか後に挿入するかの別）

を登録でき、pTeX は文章を読み込むたびにその文字がテーブルに登録されているかどうかを調べ、登録されていればそのペナルティを文字の前後適切な位置に挿入する。

禁則テーブルに情報を登録する手段として、以下のプリミティブが追加されている。

▶ `\prebreakpenalty <16-bit number>=<number>`

指定した文字の前方にペナルティを挿入する。正の値を与えると行頭禁則の指定にあたる。例えば `\prebreakpenalty`.=10000` とすれば、句点の直前に 10000 のペナルティが付けられ、行頭禁則文字の対象となる。

▶ `\postbreakpenalty <16-bit number>=<number>`

指定した文字の後方にペナルティを挿入する。正の値を与えると行末禁則の指定にあたる。例えば `\postbreakpenalty` (=10000` とすれば、始め丸括弧の直後に 10000 のペナルティが付けられ、行末禁則文字の対象となる。

和文文字、欧文文字の区別無しに指定できる。また、同一の文字に対して両方のペナルティを同時に与えるような指定はできない（もし両方指定された場合、後から指定されたものに置き換えられる）。“{}” によるグルーピング規則が適用される。禁則テーブルには 256 文字分の領域しかないので、禁則ペナルティを指定できる文字数は最大で 256 文字までである。禁則テーブルからの登録の削除はペナルティ値 0 をグローバルに設定することによって行われる^{*2}。

^{*1} ペナルティとは、行分割時やページ分割時に「その箇所がブレイクポイントとしてどの程度適切であるか」を示す一般的な評価値である（適切であれば負値、不適切であれば正值とする）。絶対値が 10000 以上のペナルティは無限大として扱われる。

^{*2} 0 をローカルに設定した場合でも、その設定が最も外側のグループ（ε-TeX 拡張でいう `\currentgrouplevel` が 0 のとき）には削除される。「ローカルで削除された場所に別の禁則ペナルティ設定がグローバルで行われ

► `\jcharwidowpenalty=<number>`

パラグラフの最終行が「す。」のように孤立するのを防ぐためのペナルティを設定する。

2 文字間のスペース

► `\kanjiskip=<skip>`

連続する和文文字間に標準で入るグルーを設定する。段落途中でこの値を変えても影響はなく、段落終了時の値が段落全体にわたって用いられる。



和文文字を表すノードが連続した場合、その間に `\kanjiskip` があるものとして行分割やボックスの寸法計算が行われる。`\kanjiskip` の大部分はこのように暗黙のうちに挿入されるものであるので、`\lastskip` などで取得することはできない。

その一方で、ノードの形で明示的に挿入される `\kanjiskip` も存在する。このようになるのは次の場合である：

- 水平ボックス (`\hbox`) が和文文字で開始しており、そのボックスの直前が和文文字であった場合、ボックスの直前に `\kanjiskip` が挿入される。
- 水平ボックス (`\hbox`) が和文文字で終了しており、そのボックスの直後が和文文字であった場合、ボックスの直後に `\kanjiskip` が挿入される。
- 連続した和文文字の間にペナルティがあった場合、暗黙の `\kanjiskip` が挿入されないので明示的にノードが作られる。



しばしば

```
\hbox to 15zw{%  
  \kanjiskip=0pt plus 1fil      あ「 い 」う え , お  
  あ「い」うえ, お}
```

のように `\kanjiskip` に無限の伸長度を持たせることで均等割付を行おうとするコードを見かけますが、連続する和文文字の間にはメトリック由来の空白と `\kanjiskip` は同時には入らないので、上に書いたコードは不適切である。

► `\xkanjiskip`

和文文字と欧文文字の間に標準で入るグルーを設定する。段落途中でこの値を変えても影響はなく、段落終了時の値が段落全体にわたって用いられる。



`\kanjiskip` と異なり、`\xkanjiskip` はノードの形で挿入される。この挿入処理は段落の行分割処理の直前や、`\hbox` を閉じるときに行われるので、「どこに `\xkanjiskip` が入っているか」を知るためには現在の段落や `\hbox` を終了させる必要がある。

► `\xspcode<8-bit number>=<0-3>`

コード番号が `<8-bit number>` の欧文文字の周囲に `\xkanjiskip` が挿入可能が否かを 0-3 の値で指定する。それぞれの意味は次の通り：

- 0 欧文文字の前側、後側ともに `\xkanjiskip` の挿入を禁止する。
- 1 欧文文字の前側にのみ `\xkanjiskip` の挿入を許可する。後側は禁止。
- 2 欧文文字の後側にのみ `\xkanjiskip` の挿入を許可する。前側は禁止。
- 3 欧文文字の前側、後側ともに `\xkanjiskip` の挿入を許可する。

る」ことがおこらないようにするため、最外グループ以外での 0 設定ではテーブルから削除されない。

pTeX の標準値は、数字 0-9 と英文字 A-Z, a-z に対する値は 3 (両側許可)、その他の文字に対しては 0 (両側禁止)。

► `\inhibitxspcode<16-bit number>=<0-3>`

コード番号が `<16-bit number>` の和文文字の周囲に `\xkanjiskip` が挿入可能が否かを 0-3 の値で指定する。それぞれの意味は次の通り：

- 0 和文文字の前側、後側ともに `\xkanjiskip` の挿入を禁止する。
- 1 和文文字の後側にのみ `\xkanjiskip` の挿入を許可する。前側は禁止。
- 2 和文文字の前側にのみ `\xkanjiskip` の挿入を許可する。後側は禁止。
- 3 和文文字の前側、後側ともに `\xkanjiskip` の挿入を許可する。

この `\inhibitxspcode` の設定値の情報は 256 文字分のテーブルに格納されている、未登録時は 3 (両側許可) であるとみなされる。

禁則ペナルティと同様に、デフォルト値の 3 をグローバルに設定するとテーブルからの登録が削除される。



`\xspmode` と `\inhibitxspmode` では、一見すると設定値 1 と 2 の意味が反対のように感じるかもしれない。しかし、実は両者とも

- 1: 「和文文字→欧文文字」の場合のみ許可。「欧文文字→和文文字」の場合は禁止。
- 2: 「欧文文字→和文文字」の場合のみ許可。「和文文字→欧文文字」の場合は禁止。

となっている。

► `\autospacing, \noautospacing`

連続する和文文字間に `\kanjiskip` で指定されただけのグルーを挿入する (`\autospacing`) か挿入しない (`\noautospacing`) を設定する。段落途中でこの値を変えても影響はなく、段落終了時の値が段落全体にわたって用いられる。

► `\autoxspacing, \noautoxspacing`

和文文字と欧文文字の間に `\xkanjiskip` で指定されただけのグルーを挿入する (`\autoxspacing`) か挿入しない (`\noautoxspacing`) を設定する。段落途中でこの値を変えても影響はなく、段落終了時の値が段落全体にわたって用いられる。

どちらの設定も標準では有効 (`\autospacing, \autoxspacing`) である。



すでに述べたように、`\kanjiskip` の一部と `\xkanjiskip` はノードの形で挿入される。`\noautospacing` や `\noautoxspacing` を指定しても、このノードの形で挿入自体は行われる (ただノードが `\kanjiskip` や `\xkanjiskip` の代わりに長さ 0 のグルーを表すだけ)。

これにより、例えば `\noautoxspacing` 状況下で「あ a」と入力しても間に長さ 0 のグルーがあるため「あ」と「a」の間に改行可能となることに注意。



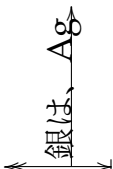
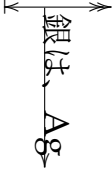
► `\showmode`

`\kanjiskip` の挿入や `\xkanjiskip` の挿入が有効になっているか否かを

- > auto spacing mode;
- > no auto xspacing mode.

という形式 (上の例では `\autospacing` かつ `\noautoxspacing` の状況) で端末やログに表示する。

表 1 pTeX のサポートする組方向

	横組	縦組	DtoU 方向	縦数式ディレクション
命令	<code>\yoko</code>	<code>\tate</code>	<code>\dtou</code>	—
字送り方向	水平右向き (→)	垂直下向き (↓)	垂直上向き (↑)	垂直下向き (↓)
行送り方向	垂直下向き (↓)	水平左向き (←)	水平右向き (→)	水平左向き (←)
使用する和文 フォント	横組用 (<code>\jfont</code>)	縦組用 (<code>\tfont</code>)	横組用 (<code>\jfont</code>) の 90° 回転	
組版例				

▶ `\inhibitglue`

この命令が実行された位置において、メトリック由来の空白の挿入を禁止する。以下の 2 点に注意。

- メトリック由来の空白が挿入されないだけであり、その代わりに `\kanjiskip` や `\xkanjiskip` が挿入されることは禁止していない。
- 本命令は現在が（非限定，限定問わず）水平モードのときしか効力を発揮しない。段落が和文文字「【」で始まり，その文字の直前にメトリック由来の空白が入ることを抑止したい場合は，次のように一旦段落を開始してから `\inhibitglue` を実行する必要がある。

```
\leavevmode\inhibitglue 【
```

3 組方向

従来の TeX では，字送り方向が水平右向き (→)，行送り方向が垂直下向き (↓) に固定されていた。pTeX では，TeX の状態として“組方向（ディレクション）”を考え，ディレクションによって字送り方向と行送り方向を変えることにしてある。なお，行は水平ボックス (horizontal box)，ページは垂直ボックス (vertical box) であるという点は，pTeX でも従来の TeX と同様である。

pTeX のサポートする組方向は横組，縦組，そして DtoU 方向^{*3}の 3 つである。また数式モード中で作られたボックスは数式ディレクションというまた別の状態になる。横組での数式ディレクション（横数式ディレクション）と DtoU 方向での数式ディレクションはそれぞれ非数式の場合と区別はないが，縦組での数式ディレクション（縦数式ディレクション）では横組を時計回りに 90 度回転させたような状態となる。従って，pTeX では縦数式ディレクションまで含めると合計 4 種類の組方向がサポートされているといえる（表 1）。

以下が，組方向の変更や現在の組方向判定に関わるプリミティブの一覧である。

^{*3} 下から上 (Down to Up) であろう。

► `\tate`, `\yoko`, `\dtou`

組方向をそれぞれ縦組, 横組, DtoU 方向に変更する. カレントの和文フォントは縦組では縦組用フォント, 横組および DtoU 方向では横組用フォントになる.

組方向の変更は, 原則として作成中のリストやボックスに何のノードも作られていない状態でのみ許される. より詳細には,

- 制限水平モード (`\hbox`), 内部垂直モード (`\vbox`) では, 上記に述べた原則通り.

`\hbox{\hsize=20em\tate ……}`

のように, ノードが作られない代入文などは組方向変更前に実行しても良い.

違反すると次のようなエラーが出る.

! Use `\tate` at top of list.

- 非制限水平モード (行分割される段落) や, 数式モード (文中数式, ディスプレイ数式問わず) での実行は禁止.
- 外部垂直モードの場合は, 次の 2 点が同時に満たされる場合のみ実行可能である.
 - `current page` の中身にはボックス, 罫線 (`\hrule`), `insertion` (`\insert`) はない. 言い換えれば, `current page` の中身はあってもマーク (`\mark`) か `whatsit` のみ^{*4}.
 - `recent contribution` の中身にもボックス, 罫線, `insertion` はない.

違反すると次のようなエラーが出る.

! Use `\tate` at top of the page.

また, ボックスを `\copy` したり中身を `\unhbox` や `\unvbox` で取り出したりする場合は, 同じ組方向のボックス内でなければならない^{*5}. 違反した場合,

! Incompatible direction list can't be unboxed.

なるエラーが出る.



`\discretionary` 命令では `\discretionary{<pre>}{<post>}{<nobreak>}` と 3 つの引数を指定するが, この 3 引数の中で組方向を変更することはできない (常に周囲の組方向が使われる).

► `\iftdir`, `\ifydir`, `\ifddir`, `\ifmdir`

現在の組方向を判定する. `\iftdir`, `\ifydir`, `\ifddir` はそれぞれ縦組, 横組, DtoU 方向であるかどうかを判定する (数式ディレクションであるかは問わない). 一方, `\ifmdir` は数式ディレクションであるかどうかを判定する.

従って, 表 1 に示した 4 つの状況のどれに属するかは以下のようにして判定できることになる.

`\iftdir`

`\ifmdir`

(縦数式ディレクション)

`\else`

(通常の縦組)

^{*4} `current page` の中身にグルー, カーン, ペナルティがあるのは, それらの前にボックス, 罫線, `insertion` が存在する場合にのみである.

^{*5} 数式ディレクションか否かは異なっても良い.

```

\fi
\else\ifysdir
    (横組)
\else
    (DtoU 方向)
\fi

```

► `\iftbox <number>`, `\ifybox <number>`, `\ifdbbox <number>`, `\ifmbox <number>`

`<number>` 番のボックスの組方向を判定する。`<number>` は有効な `box` レジスタでなければならない。

バージョン p3.7 までの pTeX では、ボックスが一旦ノードとして組まれてしまうと、通常の縦組で組まれているのか、それとも縦数式ディレクションで組まれているのかという情報が失われていた。しかし、それでは後述の「ベースライン補正の戻し量」を誤り、欧文の垂直位置が揃わないという問題が生じた ([1])。この問題を解決する副産物として、バージョン p3.7.1 で `\ifmbox` プリミティブが実装された。

4 ベースライン補正

```

\tbaselineshift \ybaselineshift
\textbaselineshiftfactor \scriptbaselineshiftfactor \scriptscriptbaselineshiftfactor

```

5 和文フォント

► `\jfam=<number>`

現在の和文数式フォントファミリの番号を格納する^{*6}。現在の欧文数式ファミリの番号を格納する `\fam` と原理的に同じ番号を指定することは原理的には可能だが、数式ファミリは和文・欧文共用であるので実際には異なる値を指定することになる。

欧文フォントが設定されている数式ファミリの番号を `\jfam` に指定し、数式中で和文文字を記述すると

! Not two-byte family.

というエラーが発生する。逆に、和文フォントが設定されているファミリの番号を `\fam` に指定し、数式中に欧文文字を記述すると

! Not one-byte family.

というエラーが発生する。

► `\jfont`, `\tfont`

欧文フォントを定義したり、現在の欧文フォントを取得する `\font` の和文版である。

^{*6} pTeX, upTeX では 0-15 の範囲が許される。ε-pTeX, ε-upTeX では欧文の `\fam` と共に 0-255 に範囲が拡張されている。

一応 `\jfont` が「和文の横組用フォント」の、`\tfont` が「和文の縦組用フォント」のために用いる命令である。

- フォントを定義する際は、欧文フォント・和文の横組用フォント・和文の縦組用のフォントのいずれも `\font`, `\jfont`, `\tfont` のどれを用いても定義できる。
- `\the` 等で「現在のフォント」を取得する際には、`\jfont` で「和文の横組用フォント」を、`\tfont` で「和文の縦組用フォント」を返す。

► `\kcatcode⟨16-bit number⟩=⟨16–18⟩`

pTeX では、和文文字に 16 (*kanji*), 17 (*kana*), 18 (*other_kchar*) のいずれかのカテゴリコードを割り当てている。これらの値による動作の違いは次のようになる：

- TeX82 では、「複数文字からなる命令」(control word) にはカテゴリコードが 11 (*letter*) の文字しか使用できないことになっていたが、pTeX では 16, 17 の和文文字も合わせて使用することが出来る。
カテゴリコードが 18 の和文文字は control word 中には使用できない^{*7}。
- 前で説明した `\jcharwidowpenalty` は、カテゴリコードが 16, 17 の和文文字の前にのみ挿入されうるもので、カテゴリコードが 18 の和文文字の前には挿入されない。

`\kcatcode` は DVI 中の上位バイト（すなわち、JIS コードでいう区ごと）に値が設定可能である。初期状態では 1, 2, 7–15, 85–94 区の文字が 18, 3–6 区の文字が 17, 16–84 区の文字が 16 である。



upTeX においては、`\kcatcode` は大きく仕様変更されている。

6 文字コード変換，漢数字

pTeX の内部コードは環境によって異なる。そこで、異なる文字コード間で同じ文字を表現するため、文字コードから内部コードへの変換プリミティブが用意されている。また、漢数字を出力するプリミティブも用意されている^{*8}。

► `\kuten ⟨16-bit number⟩`

区点コードから内部コードへの変換を行う。16 進 4 桁の上 2 桁が区，下 2 桁が点であると解釈する。たとえば、`\char\kuten"253C` は、「怒」(37 区 60 点) である。

► `\jis ⟨16-bit number⟩`, `\euc ⟨16-bit number⟩`, `\sjis ⟨16-bit number⟩`

それぞれ JIS コード，EUC コード，Shift_JIS コードから内部コードへの変換を行う。たとえば、`\char\jis"346E`, `\char\euc"B0A5`, `\char\sjis"8A79` は、それぞれ「喜」，「哀」，「楽」である。

^{*7} `\` のように一文字命令 (control symbol) としての使用は可能である。

^{*8} 実は `\kansuji`, `\kansujichar` プリミティブは p3.1.1 でいったん削除され、p3.1.2 で復活したという経緯がある。

► `\kansuji <number>, \kansujichar <0-9>=<16-bit number>`

`\kansuji` は、続く数値 `<number>` を漢数字の文字列で出力する。出力される文字は `\kansujichar` で指定できる（デフォルトは「〇一二三四五六七八九」）。たとえば

```
\kansuji 1978 年
```

は「一九七八年」と出力され、

```
\kansujichar1=`壺
```

```
\kansujichar2=\euc"C6F5\relax
```

```
\kansujichar3=\jis"3B32\relax
```

```
\kansuji 1234
```

は「壺貳参四」と出力される。

以上に挙げたプリミティブ (`\kuten`, `\jis`, `\euc`, `\sjis`, `\kansuji`) は内部整数を引数にとるが、実行結果は文字列であることに注意。また、`pTeX` は JIS X 0213 には対応せず、JIS X 0208 の範囲のみ扱える。

```
\newcount\hoge
```

```
\hoge="2423
```

9251, 九二五一

```
\the\hoge, \kansuji\hoge\
```

42147, い

```
\jis\hoge, \char\jis\hoge\
```

一七〇一

```
\kansuji1701
```

7 長さ単位

`pTeX` では `TeX82` に加えて以下の単位が使用可能である：

► **zw**

現在の和文フォント（通常の縦組のときは縦組用フォント、それ以外のときは横組用フォント）における「全角幅」。例えばこの文書の本文では $1\text{zw} = 10.12534\text{pt}$ である。

► **zh**

現在の和文フォント（通常の縦組のときは縦組用フォント、それ以外のときは横組用フォント）における「全角高さ」。例えばこの文書の本文では $1\text{zh} = 9.64365\text{pt}$ である。



より正確に言えば、**zw**, **zh** はそれぞれ標準の文字クラス（文字クラス 0）に属する和文文字の幅、高さとの和を表す。

ただ、`pTeX` の標準和文フォントメトリックの一つである `min10.tfm` では、 $1\text{zw} = 9.62216\text{pt}$, $1\text{zh} = 9.16443\text{pt}$ となっており、両者の値は一致していない。JIS フォントメトリックでも同様の寸法となっている。一方、実際の表示に使われる和文フォントの多くは、 $1\text{zw} = 1\text{zh}$, すなわち正方形のボディに収まるようにデザインされているから、これと合致しない。したがって、単位 **zh** はあまり意味のある値とはいえない。

なお、`japanese-otf` (OTF パッケージ) が用いているフォントメトリックは $1\text{zw} = 1\text{zh}$ である。

► **Q, H**

両者とも 0.25mm ($7227/10160\text{sp}$) を意味する。写植機における文字の大きさの単位である **Q** 数（級数）と、字送り量や行送り量の単位である歯数に由来する。

参考文献

- [1] aminophen, 「縦数式ディレクションとベースライン補正」, 2016/09/05,
<https://github.com/texjporg/platex/issues/22>

索引

Symbols			
\autospaceing	4	\kcatcode	8
\autoxspacing	4	\kuten	8
\dtou	6	\noautospaceing	4
\euc	8	\noautoxspacing	4
\ifdbox	7	\postbreakpenalty	2
\ifddir	6	\prebreakpenalty	2
\ifmbox	7	\showmode	4
\ifmdir	6	\sjis	8
\iftbox	7	\tate	6
\iftdir	6	\tfont	7
\ifybox	7	\xkanjiskip	3
\ifydir	6	\xspcode	3
\inhibitglue	5	\yoko	6
\inhibitxspcode	4		
\jcharwidowpenalty	3	H	
\jfam	7	H	9
\jfont	7		
\jis	8	Q	
\kanjiskip	3	Q	9
\kansuji	9		
\kansujichar	9	Z	
		zh	9
		zw	9