

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹.

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: *Revised Proposal to Encode Mathematical Diagonals*

2. Requester's name: *Deborah Anderson (University of California, Berkeley, Script Encoding Initiative), Barbara Beeton (American Mathematical Society), Laurentiu Iancu, Murray Sargent (Microsoft Corporation)*

3. Requester type (Member body/Liaison/Individual contribution): *SEL liaison contribution*

4. Submission date: *2010-January-12*

5. Requester's reference (if applicable): *Replaces INCITS/L2 document L2/09-397*

6. Choose one of the following:
This is a complete proposal: *Yes*
(or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
a. This proposal is for a new script (set of characters): *No*
Proposed name of script: _____
b. The proposal is for addition of character(s) to an existing block: *Yes*
Name of the existing block: *Miscellaneous Mathematical Symbols-A*

2. Number of characters in proposal: *2*

3. Proposed category (select one from below - see section 2.2 of P&P document):
A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection)
C-Major extinct D-Attested extinct E-Minor extinct
F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*
b. Are the character shapes attached in a legible form suitable for review? *Yes*

5. Fonts related:
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? *The authors*
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): *The authors*

6. References:
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*

7. Special encoding issues:
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*
Suggested UCD character properties in Section 4

8. Additional Information:
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see <http://www.unicode.org/Public/UNIDATA/UCD.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3702-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	No
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents:	<i>David Carlisle (Numerical Algorithms Group), members of the UTC</i> <i>Examples in proposal</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	No
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>Common</i> <i>Supporting evidence in Section 5</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>Yes</i> <i>Supporting evidence in Section 5</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	No
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	No
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>Yes</i> <i>Yes</i> <i>Disunification explained in Section 2</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	No
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	No
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference:	No

1. Introduction

This proposal is to encode two mathematical diagonal bar symbols, a rising diagonal and a falling diagonal. The two symbols can be found in contemporary use in mathematical publications, for instance in the notation of spaces of double cosets. The proposed symbols differ from other solidi by their slant. They are proposed to be disunified from the box-drawing diagonals, U+2571 BOX DRAWING LIGHT DIAGONAL UPPER RIGHT TO LOWER LEFT and U+2572 BOX DRAWING LIGHT DIAGONAL UPPER LEFT TO LOWER RIGHT, whose slant may vary between fonts. As the evidence samples will demonstrate, the slant is significant, the symbols being used contrastively with other mathematical solidi, including U+2215 DIVISION SLASH, U+29F5 REVERSE SOLIDUS OPERATOR, and U+2216 SET MINUS.

The proposed symbols are part of the STIX PUA collection, whose standardization in Unicode is an ongoing project. A description of the STIX Fonts Project [1] and the STIX PUA set can be found in the INCITS/L2 proposal documents [L2/09-261](#) and [L2/09-262](#) [2].

2. Disunification

The proposed diagonals correspond to the LaTeX entities `\diagup` and `\diagdown`. In early versions of Unicode, the box-drawing symbols U+2571 BOX DRAWING LIGHT DIAGONAL UPPER RIGHT TO LOWER LEFT and U+2572 BOX DRAWING LIGHT DIAGONAL UPPER LEFT TO LOWER RIGHT had representative glyphs drawn at a 45°/135° angle. This implied a unification of the Unicode characters to use for `\diagup` and `\diagdown` with the box-drawing diagonals. However, as the Standard evolved, the representative glyphs of the box-drawing diagonals changed slope. Fonts also adopted glyphs of varying slope, making the box-drawing diagonals inadequate in contexts where a specific slant is required, such as in mathematical notation. This document proposes the disunification of the mathematical diagonals from the box-drawing diagonals by encoding two distinct characters to represent `\diagup` and `\diagdown` in Unicode.

The proposed characters differ from existing solidi by their slope. Figure 1 reproduces several solidi as they appear in the Unicode 5.2 code charts. In contrast, the proposed symbols are drawn at a 45°/135° angle with the baseline, which is not allowed to change as glyphic variation.

2215	/	DIVISION SLASH • generic division operator → 002F / solidus → 2044 / fraction slash	29F5	\	REVERSE SOLIDUS OPERATOR → 005C \ reverse solidus → 2216 \ set minus
2216	\	SET MINUS → 005C \ reverse solidus	29F6	⌘	SOLIDUS WITH OVERBAR
			29F7	↘	REVERSE SOLIDUS WITH HORI → 2340 ↘ apl functional symb
			Large operators		
2571	/	BOX DRAWINGS LIGHT DIAGONAL UPPER RIGHT TO LOWER LEFT	29F8	/	BIG SOLIDUS → 2215 / division slash
2572	\	BOX DRAWINGS LIGHT DIAGONAL UPPER LE TO LOWER RIGHT	29F9	\	BIG REVERSE SOLIDUS = z notation schema hiding → 2216 \ set minus

Figure 1: Solidi present in the Unicode 5.2 code charts.

In particular, the box-drawing diagonals, U+2571 BOX DRAWING LIGHT DIAGONAL UPPER RIGHT TO LOWER LEFT and U+2572 BOX DRAWING LIGHT DIAGONAL UPPER LEFT TO LOWER RIGHT, have representative glyphs that changed between versions of Unicode, to align with other box-drawing symbols, whether those formed square or rectangular box cells. In Unicode 3.0, the diagonals aligned with square box pieces, whereas later versions adopted a rectangular style. Figure 2 illustrates the glyph changes between Unicode versions, in contrast with the glyphs of the proposed characters.

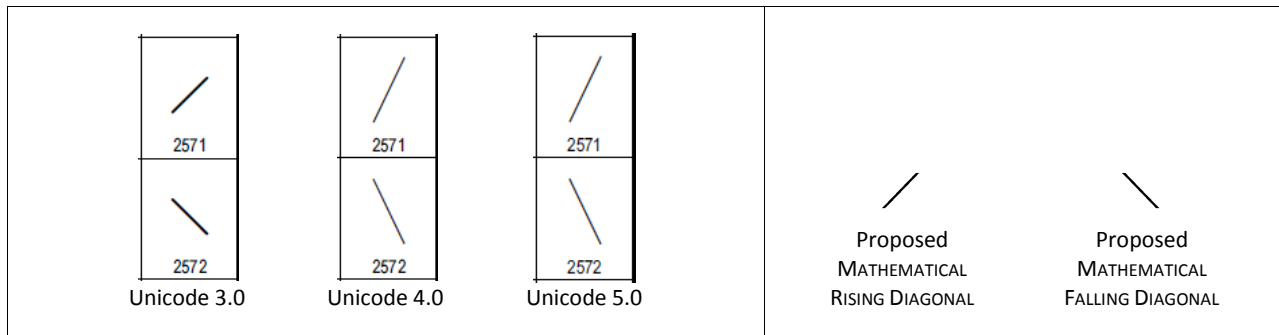


Figure 2: (a) Evolution of the representative glyphs for the box-drawing symbols U+2571 and U+2572 in the Unicode code charts [3] (left). (b) Glyphs of the proposed rising and falling diagonals (right).

Fonts too commonly scale the box-drawing symbols to their preferred aspect ratio. As a result, U+2571 and U+2572 have glyphs of different slope in different fonts, which makes them inadequate for contexts where a certain fixed slope is required. Figure 3 compares the glyphs of several solidi in a few Microsoft Office and Microsoft Windows 7 fonts.

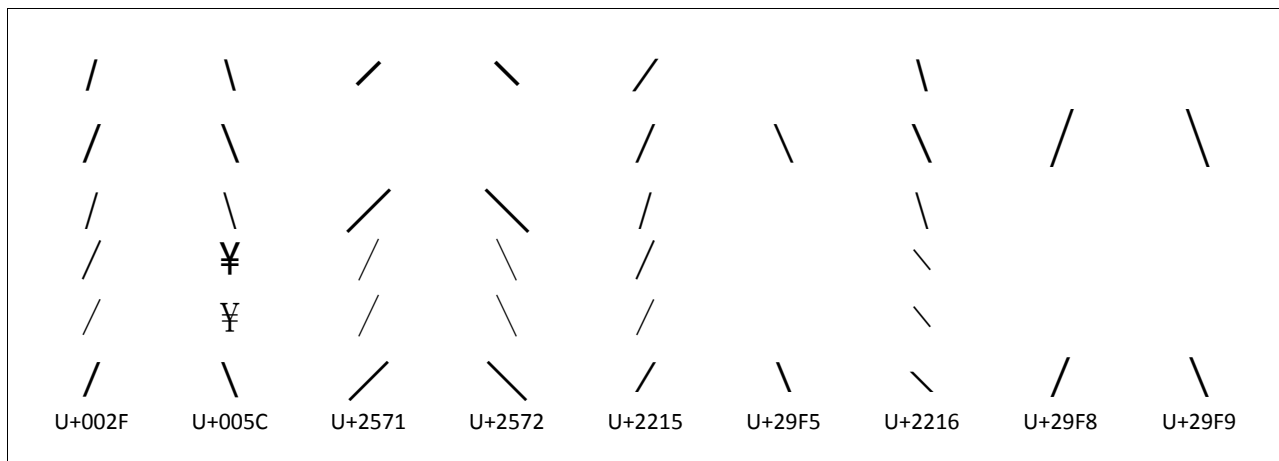


Figure 3: Glyphs of existing solidi in a few Microsoft Office and Microsoft Windows 7 fonts at size 16: Arial Unicode MS, Cambria Math, Lucida Sans Unicode, MS Gothic, MS Mincho, and Segoe UI Symbol. Note, in particular, the varying slopes of the box-drawing symbols U+2571 and U+2572.

3. Mappings

The following public entity sets and symbol lists were consulted in search for mappings to Unicode characters for representing the proposed symbols:

- SGML Entity Set ISO/IEC DTR 9573-13:1990 – <http://www.dcarlisle.demon.co.uk/9573/>;
- XML Entity Definitions for Characters – latest draft <http://www.w3.org/2003/entities/2007doc/>; and data files <http://www.w3.org/2003/entities/2007xml/>;
- MathML – latest MathML 3.0 draft <http://www.w3.org/TR/MathML3/>;
- STIX Fonts Project [1] and PUA collection [2] – the proposed symbols are mapped to code points in the PUA;
- The Comprehensive LaTeX Symbol List [4];
- Elsevier Science Grid in Unicode – <http://info.sciencedirect.com/techsupport/xmlsgml/dtd50/esgrid.pdf>;
- OASIS DocBook – <http://www.oasis-open.org/docbook/specs/wd-docbook-xmlcharent-0.3.pdf>;
- Adobe Systems Inc. Mathematical Pi Fonts One–Six;
- Design Science Font Encoding Tables and MTCODE Encoding Tables – http://www.dessci.com/en/support/mathtype/tech/encodings/font_enc.htm and <http://www.dessci.com/en/support/mathtype/tech/encodings/mtcode.htm>;
- Wolfram Research, Inc., Mathematica Listing of Named Characters – <http://reference.wolfram.com/mathematica/guide/ListingOfNamedCharacters.html> and <http://www.mathmlcentral.com/characters/>.

The “XML Entity Definitions for Characters” file <http://www.w3.org/2003/entities/2007xml/unicode.xml> (retrieved January 2010) maps LaTeX entities `\diagup` and `\diagdown` to U+2571 and U+2572, respectively. However, LaTeX entities are typically used by name rather than via mappings to Unicode code points. The editor of the file was notified about the proposed symbols and their mapping.

4. Proposed Characters

The proposed characters are

- U+27CB MATHEMATICAL RISING DIAGONAL
- U+27CD MATHEMATICAL FALLING DIAGONAL

allocated in the remaining unassigned code points of the Miscellaneous Mathematical Symbols-A block. Their glyphs must form a 45°/135° angle with the baseline, as shown earlier in Figure 2b.

The proposed characters have the properties suggested below, similar to those of other mathematical symbols, such as U+2215 DIVISION SLASH, U+29F5 REVERSE SOLIDUS OPERATOR, U+29F8 BIG SOLIDUS, and U+29F9 BIG REVERSE SOLIDUS. The two symbols form a mirroring pair.

UnicodeData.txt entry (gc, ccc, bc, dt, Bidi_M, etc.)	bmg	Script	lb	Math
27CB;MATHEMATICAL RISING DIAGONAL;Sm;0;ON;;;;;Y;;;;;	27CD	Common	AL	Yes
27CD;MATHEMATICAL FALLING DIAGONAL;Sm;0;ON;;;;;Y;;;;;	27CB	Common	AL	Yes

Table 1: UCD properties of the proposed characters (Math = Yes derives implicitly from gc = Sm).

5. Supporting Evidence

This section attests the proposed symbols with from samples from two main sources:

- Several LaTeX symbol charts, consolidated in the “The Comprehensive LaTeX Symbol List” [4];
- An article [5] published by the American Mathematical Society, which illustrates the contrastive use of the proposed diagonals together with the other solidi.

TABLE 185: Miscellaneous \mathcal{AMS} Math Symbols					
\sphericalangle	<code>\angle</code>	\blacktriangledown	<code>\blacktriangledown</code>	\mho	<code>\mho</code>
\backprime	<code>\backprime</code>	\diagdown	<code>\diagdown</code>	\sphericalangle	<code>\sphericalangle</code>
\bigstar	<code>\bigstar</code>	\diagup	<code>\diagup</code>	\square	<code>\square</code>
\blacklozenge	<code>\blacklozenge</code>	\eth	<code>\eth</code>	\triangledown	<code>\triangledown</code>

TABLE 188: Miscellaneous mathabx Math Symbols					
\circ	<code>\degree</code>	\fourth	<code>\fourth</code>	\measuredangle	<code>\measuredangle</code>
\diagdown	<code>\diagdown</code>	$\#$	<code>\hash</code>	\pitchfork	<code>\pitchfork</code>
\diagup	<code>\diagup</code>	∞	<code>\infty</code>	\propto	<code>\propto</code>
\oslash	<code>\diameter</code>	\leftthreetimes	<code>\leftthreetimes</code>	\rightthreetimes	<code>\rightthreetimes</code>

MnSymbol additionally defines synonyms for some of the preceding symbols:

\dashv	<code>\dashv</code>	(same as <code>\leftvdash</code>)
\diagdown	<code>\diagdown</code>	(same as <code>\nwseline</code>)
\diagup	<code>\diagup</code>	(same as <code>\neswline</code>)

Figure 4: Samples illustrating the entities “`\diagup`” and “`\diagdown`” in several LaTeX packages, consolidated in “The Comprehensive LaTeX Symbol List” [4] (one package defining additional synonyms). From the top: “`AMSsymbols`” package, p. 67; “`mathabx`” package, p. 68; “`MnSymbol`” package, p. 34.

MnSymbol additionally defines synonyms for some of the preceding symbols:

\ndashv	<code>\ndashv</code>	(same as <code>\nleftvdash</code>)
\ndiagdown	<code>\ndiagdown</code>	(same as <code>\nnwseline</code>)
\ndiagup	<code>\ndiagup</code>	(same as <code>\nneswline</code>)

Figure 5: Samples illustrating the negated entities “`\ndiagup`” and “`\ndiagdown`” in the LaTeX package “`MnSymbol`,” from p. 35 of [4]. Combining character sequences employing U+0338 COMBINING LONG SOLIDUS OVERLAY and U+20E5 COMBINING REVERSE SOLIDUS OVERLAY (<U+27CB, U+20E5> and <U+27CD, U+0338>) can be used to represent the negated symbols instead of proposing precomposed characters.

over the set of double cosets space $H \backslash G / B$ and that any isotopic component is of the form $\text{ind}_{B(x)}^H \chi_{\ell(x)}$, where $B(x) = H \cap \psi(x)B\psi(x)^{-1}$, $x \in H \backslash G / B$, $\ell(x) = \text{Ad}^*(\psi(x))\ell|_{\mathfrak{h}}$ and $\psi : H \backslash G / B \rightarrow G$ is a section for the double cosets space. So,

Figure 6: Excerpt from p. 819 of [5] illustrating the proposed diagonals inline text.

$$\mathcal{I}^{\mathfrak{g}/\mathfrak{c}} = \mathcal{I}^{\mathfrak{g}/\mathfrak{b}} \setminus \{l\} = \{k_1 < \dots < k_{p-1}\},$$

whence, by intersecting with $\mathcal{I}^{\mathfrak{g}/\mathfrak{h}}$, we obtain that $\mathcal{I}^{\mathfrak{b},\mathfrak{c}} = \mathcal{I}^{\mathfrak{b},\mathfrak{b}} \setminus \{l\}$.

We encounter the following two cases which must be discussed also in the proof of Lemma 3.9.

a): $Z_l \notin \mathfrak{h} + \mathfrak{g}_{l+1}$, which means that $l \in \mathcal{I}^{\mathfrak{g}/\mathfrak{h}}$ and so $\mathcal{I}^{\mathfrak{b},\mathfrak{b}} = \mathcal{I}^{\mathfrak{b},\mathfrak{c}} \cup \{l\}$.

b): $Z_l \in \mathfrak{h} + \mathfrak{g}_{l+1}$, whence $Z_l \in \mathfrak{h}$ while we assume (2.3). In this case we have $l \notin \mathcal{I}^{\mathfrak{g}/\mathfrak{h}}$ and so $\mathcal{I}^{\mathfrak{b},\mathfrak{b}} = \mathcal{I}^{\mathfrak{b},\mathfrak{c}}$.

Since $\mathfrak{g}_l \subset \mathfrak{g}_{i+1}$ for all $i < l$, it follows that

$$\mathfrak{h} + \text{Ad}(g)\mathfrak{c} + \mathfrak{g}_{i+1} = \mathfrak{h} + \text{Ad}(g)(\mathfrak{b} + \mathfrak{g}_l) + \mathfrak{g}_{i+1} = \mathfrak{h} + \text{Ad}(g)\mathfrak{b} + \mathfrak{g}_{i+1}, \quad \forall g \in G.$$

This tells us that $\mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{c}) = \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b}) \setminus \{l\}$.

We can also deduce from this equality that all the nonempty Zariski open subsets \mathcal{U}_i° ($i \in \mathcal{I}^{\mathfrak{b},\mathfrak{c}}$) of G , defining the canonical Zariski open subset $\mathcal{U}^\circ = \bigcap_{i \in \mathcal{I}^{\mathfrak{b},\mathfrak{c}}} \mathcal{U}_i^\circ$ as in Remark 3.4 for $H \backslash G / C$, coincide with the \mathcal{U}_i ($i \in \mathcal{I}^{\mathfrak{b},\mathfrak{b}} \setminus \{l\}$) for $H \backslash G / B$. In particular, from $\mathcal{I}^{\mathfrak{b},\mathfrak{b}} \supseteq \mathcal{I}^{\mathfrak{b},\mathfrak{c}}$, we obtain that $\mathcal{U} \subseteq \mathcal{U}^\circ$.

Hence in the case b), we have that

$$\mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b}) = \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{c}),$$

since $l \notin \mathcal{I}^{\mathfrak{g}/\mathfrak{h}} \supset \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b})$ by Proposition 3.3 (2).

In the first case, we consider the following two eventualities.

a-1) $l \in \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b})$: This case is characterized by the existence of a nonempty Zariski open subset \mathcal{U}_l of G , given by Definition 3.2 (2), such that for every $g \in \mathcal{U}_l$, we have $Z_l \notin \mathfrak{h} + \text{Ad}(g)\mathfrak{b}$ (since $\mathfrak{g}_{l+1} \subseteq \text{Ad}(g)\mathfrak{b}$ for $g \in G$).

a-2) $l \notin \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b})$: Here we have $\mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{c}) = \mathcal{I}(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}/\mathfrak{b})$ as in the case b). Hence for every g in a nonempty Zariski open subset \mathcal{U}_l of G , we have that $Z_l \in \mathfrak{h} + \text{Ad}(g)\mathfrak{b}$.

Remark 3.8. For every $(w, \tilde{g}) \in G_l \times H \backslash G / B$, we have $H \cdot w g w^{-1} \cdot B = \tilde{g}$. This is an

Figure 7: Excerpt from p. 825 of [5] illustrating the proposed diagonals being used contrastively with other solidi. The proposed diagonals are highlighted in green, forward (division) solidi (U+2215) in yellow, backward (reverse) solidi (U+29F5) in blue, and (small) set minus (U+2216) in purple.

6. Code Charts

A UniBook code chart is appended at the end of this document. UniBook source project files and a font will be made available to the editors. Updated UCD files (such as `BidiMirroring.txt` and `MathClass*.txt` to accompany UTR #25) will also be made available to the Unicode Technical Committee.

7. References

- [1] The STIX Font Project, <http://www.stixfonts.org/>, an activity of the STIX Pub companies listed at <http://www.stixfonts.org/stipubs.html>. The latest version of the STIX table is linked from <http://www.ams.org/STIX/>.
- [2] Deborah Anderson et al., "Preliminary Proposal to Encode Characters from the STIX PUA Collection," INCITS/L2 documents L2/09-261 and L2/09-262, <http://www.unicode.org/L2/L2009/09261-stix-pua-proposal-1.pdf> and <http://www.unicode.org/L2/L2009/09262-stix-pua-proposal-2.pdf>.
- [3] Internet Archive Wayback Machine, archive of the Unicode Box Drawing code chart, queried Jan. 2010, http://web.archive.org/web/*/http://www.unicode.org/charts/PDF/U2500.pdf.
- [4] Scott Pakin, "The Comprehensive LaTeX Symbol List," version of Jan. 2008, <http://www.ctan.org/tex-archive/info/symbols/comprehensive/>.
- [5] Jawhar Abdennadher and Jean Ludwig, "Disintegrating Tensor Representations of Nilpotent Lie Groups," *Transactions of the American Mathematical Society*, vol. 361, no. 2, Feb. 2009, pp. 819–848.

	27C	27D	27E
0	▨	▨	▨
1	▨	▨	▨
2	▨	▨	▨
3	▨	▨	▨
4	▨	▨	▨
5	▨	▨	▨
6	▨	▨	▨
7	▨	▨	▨
8	▨	▨	▨
9	▨	▨	▨
A	▨	▨	▨
B	▨ 27CB	▨	▨
C	▨	▨	▨
D	▨ 27CD	▨	▨
E	▨	▨	▨
F	▨	▨	▨

Miscellaneous Symbol

27CB / MATHEMATICAL RISING DIAGONAL
= \diagup
• drawn at a 45-degree angle with the horizontal
→ 2215 / division slash
→ 2571 / box drawings light diagonal upper right to lower left

Miscellaneous Symbol

27CD \ MATHEMATICAL FALLING DIAGONAL
= \diagdown
• drawn at a 135-degree angle with the horizontal
→ 2216 \ set minus
→ 2572 \ box drawings light diagonal upper left to lower right
→ 29F5 \ reverse solidus operator