Goodness-of-Fit Tests DIFF1 and DIFF2 for Locally-Normalized Supernova Spectra

David J. Jeffery\textsuperscript{1}, Wesley Ketchum\textsuperscript{1}, David Branch\textsuperscript{1}, E. Baron\textsuperscript{1}, Abouazza Elmhamdi\textsuperscript{2,3}, & I.J. Danziger\textsuperscript{3}

ABSTRACT

Two quantitative tests DIFF1 and DIFF2 for measuring goodness-of-fit between two locally-normalized supernova spectra are presented. Locally-normalized spectra are obtained by dividing a spectrum by the same spectrum smoothed over a wavelength interval relatively large compared to line features, but relatively small compared to continuum features. DIFF1 essentially measures the mean relative difference between the wave patterns of locally-normalized spectra and DIFF2 is DIFF1 minimized with respect to a relative logarithmic wavelength shift between the spectra: the shift is an artificial relative Doppler shift. Both DIFF1 and DIFF2 measure the similarity of spectra: DIFF1 puts more weight on overall physical similarity in spectrum formation; DIFF2, just on the similarity of the line patterns in the spectra because the shift compensates for some physical distinction in the supernovae. Both tests are useful in ordering supernovae into empirical groupings for further analysis. We present some examples of locally-normalized spectra for Type IIb supernova SN 1993J with some analysis of these spectra. The spectra include two \textit{HST} spectra that have not been published before. We also give an example of fitted locally-normalized spectra and, as an example of the utility of DIFF1 and DIFF2, some preliminary statistical results for hydrogen-deficient core-collapse (HDCC) supernova spectra. This paper makes use of and refers to material to found at the first author’s online supernova spectrum database SUSPEND (SUpernovae Spectra PENDing further analysis).\textsuperscript{1}


\textsuperscript{1}Homer L. Dodge Department of Physics & Astronomy, University of Oklahoma, 440 W. Brooks St., Norman, Oklahoma 73019, U.S.A.

\textsuperscript{2}ICTP- The International Centre for theoretical Physics, Strada Costiera 11 34014, Trieste, Italy

\textsuperscript{3}INAF-Osservatorio Astronomico di Trieste, via G.B. Tiepolo 11, 34131 Trieste, Italy

\textsuperscript{1}See http://www.nhn.ou.edu/~jeffery/astro/sne/spectra/spectra.html.
1. INTRODUCTION

An ordinary \( \chi^2 \)-like test measuring the goodness-of-fit of two supernova spectra often fails to be consistent with what the eye sees qualitatively: good agreement to the eye can be poor or moderate agreement by a \( \chi^2 \)-like test. Since the human eye is an excellent pattern recognition tool, judgments based on eye comparisons informed by a specialist’s understanding of spectrum formation are often preferred over \( \chi^2 \)-like tests or other formulaic tests in identifying similarity and/or goodness-of-fit between two spectra. Human judgment, of course, is qualitative and to some degree subjective. Thus, it would be advantageous to have tests that measure spectrum similarity more consistently with what the eye sees and yet be quantitative and objective. Such tests could be applied in an automated fashion and, one hopes, would show correlations missed by the eye.

In this paper, we present two quantitative tests DIFF1 and DIFF2 for measuring goodness-of-fit between two supernova spectra. The tests are both measures of the relative difference in line patterns. (The formulae for the tests are given in § 4.1 and 4.2.) Both tests depend on what we call local normalization which largely removes continuum features without distorting the line features too much. Local normalization reduces the continuum shape to an apparent flat line of height 1 by dividing the original spectrum by a version of the spectrum smoothed over a large smoothing length. (The equivalent of local normalization has frequently been used for synthetic spectrum calculations simply by inputting a flat continuum at the base of a model atmosphere or by dividing the spectrum by the continuum. In the context of synthetic supernova spectra see, e.g., Dessart & Hillier (2005).) Ordinary (or global) normalization is just to divide the spectrum by a constant.

DIFF1 and DIFF2 evolved from the DIFF test presented in Branch et al. (2006b). DIFF1 essentially measures the mean relative difference between the wave patterns of locally-normalized spectra and DIFF2 is DIFF1 minimized with respect to a relative logarithmic wavelength shift between the spectra: the shift is an artificial relative Doppler shift. From this description, the reader can roughly understand both tests before knowing the test formulae. Both DIFF1 and DIFF2 measure the similarity of spectra: DIFF1 puts more weight on overall physical similarity in spectrum formation; DIFF2, just on the similarity of the line patterns in the spectra. Both tests are useful in ordering supernovae into empirical groupings for further analysis.\(^2\)

\(^2\)While this paper was in preparation, analysis tools for supernova spectra that are in some respects similar to DIFF1 and DIFF2 with local normalization were described by Stephane Blondin and Avet Harutyunyan at the conference The Multicoloured Landscape of Compact Objects and their Explosive Origins, Cefalù, Sicily, 2006 June 11–24, URL: http://www.mporzio.astro.it/cefalu2006/.
There are several reasons for local normalization. First, the observed continuum can often be quite uncertain. Often the main reason for this is reddening that can affect, for example, the $B - V$ color and other colors (particularly toward the blue) by a significant fraction of a magnitude or more. Uncertainty in reddening is particularly a problem for core-collapse supernovae (i.e., Types II-P, II-L, IIn, IIb, Ib, Ic, and Ic hypernovae) which tend to arise in or near star-forming regions in their host galaxies. Foreground Galactic reddening can be reasonably confidently corrected for using the results of Schlegel et al. (1998) for Galactic reddening and Cardelli et al. (1989) and O'Donnell (1994) for the reddening law, but the host galaxy reddening can usually only be corrected for with great uncertainty relying on interstellar lines in the supernova spectra or, with less uncertainty depending on cases, on spectral modeling. For example, early spectral modeling of Type IIP supernovae can constrain reddening (Baron et al. 2003, 2004, 2005). Nevertheless, in spectral modeling one would much prefer to have reddening as a given rather than as a parameter to be fitted for—or not to be a consideration which is what local normalization helps toward. Core-collapse supernovae form a rather heterogeneous class, and so judging reddening by comparison is not usually practical. Type Ia supernovae (SNe Ia) are usually much less reddened as they do not preferentially arise in or near star forming regions and the homogeneity of most of them (e.g., Branch et al. 2006b) can allow reddening correction by comparison in many cases. But even SNe Ia can be highly reddened (e.g., SN 1986G (Phillips et al. 1987)) and there are peculiar SNe Ia (e.g., Branch et al. 2006b) for which comparisons to determine reddening can be uncertain.

In addition to reddening, the continuum can be in error because of errors in the broadband calibration of the spectrum. A particular problem in achieving good calibration arises because supernovae are transient, time-dependent sources: this means one usually cannot simply replace a bad observation by a good for a given epoch and this limits the ability to achieve good calibration. Another related particular problem arises from the fact that bright well-observed supernovae are relatively rare. Thus, many well-observed supernovae are from earlier instrumentation epochs. One would like to use the data from these supernovae for analysis since they are to some degree unique even though the data may have very uncertain calibration by modern standards.

Reddening and calibration errors tend to affect the spectrum continuum multiplicatively: i.e., they change flux level by a scaling factor that varies slowly with wavelength. These kind of uncertainties are ideally and nearly practically irrelevant to the process of local normalization (see § 2.1), and thus cannot much affect analysis with the locally-normalized spectra.

A kind of continuum uncertainty that is not eliminated by local normalization is the
additive uncertainty due to contamination by extraneous sources. (See § 2.1 for why this is so.) The most common contamination is from the host galaxy emission. When a supernova is at its brightest this contamination is often small, but in later phases when the supernova dims it can become significant and dominant. Host galaxy contamination can be subtracted off using spectra of the galaxy when the supernovae is absent, but it is not always clear to analyzers of supernova spectra when this has been done. Another kind of contamination is peculiar to supernovae associated with gamma-ray bursts (GRBs): such supernovae may all be what are now called hypernovae or as we prefer Type Ic hypernovae since they seem to lack conspicuous lines of hydrogen and helium. The contamination is the UVOIR (ultraviolet-optical-infrared) afterglow of the GRB itself which tends to be a line-free continuum. Correction for this kind of contamination can be done, but with some uncertainty (Matheson et al. 2003). Yet another contamination (which unfortunately also affects the line pattern) is light echoing caused by supernova light from earlier phases reflected off dust clouds and added to the light of the current supernova phase. Light echoes, if recognized, can be corrected for, but with some error, of course. For the developments in this paper, we assume that contamination of all kinds can be adequately corrected for.

Because of all the effects mentioned above, the continuum shape of supernova spectra may be incorrect and misleading about the degree of similarity between spectrum pairs. On the other hand the line pattern in the spectra should be better for determining similarities since this is less affected by broadband uncertainties in reddening and calibration. For example, if a spectrum pair had an identical line pattern, but different continuum shapes, one would have good reason to believe that the supernovae are, in fact, highly similar and at a similar phase in their evolution and that a discrepancy in continuum shape was just caused by reddening correction and/or calibration error. In the contrasting case, where the line patterns are very different, but the continuum shapes are identical, one would conclude the supernovae are not very similar or at different phases in their evolution despite the similar continua. In both extreme cases, it is the line patterns that provide the decisive evidence, not the continua. Thus, if one can adequately correct or neglect contamination, analysis with locally-normalized spectra should give good insight into the intrinsic similarities and differences among supernovae.

A second reason for using local normalization applies even when the continuum shapes are assumed to be well known. If one is searching just for similarity of line patterns as a clue to physical similarities, then varying continuum shapes tend to obscure the similarities of the patterns both to the eye and formulaic tests. Being able to measure similarity of line patterns for heterogeneous samples of supernovae from varying phases is important in studying the time-varying supernova structure. Especially for core-collapse supernovae in which even members of the same type show considerably individuality and in which the time
coverage of their evolution can often be very incomplete, local normalization could become an important tool in analysis.

A third reason, for local normalization is that spectra are frequently analyzed using synthetic spectra calculated from highly simplified radiative transfer: e.g., analyses done using the parameterized code SYNOW (e.g., Branch et al. 2003, 2005, and references therein) Such simplified radiative transfer does not treat continuum radiative transfer with high physical realism and in fits of synthetic to observed spectra mismatches in the continuum in some regions are obvious and are not considered very important in the important results derived from the synthetic spectrum analysis: the important results being line identification and ejecta velocity structure. Local normalization of both observed and synthetic spectra could be used to reduced the distraction of mismatching continuum shapes.

In § 2 of this paper, we discuss the theoretical basis for local normalization and how we actually carry out local normalization. Section 3 gives examples of locally-normalized spectra for the Type IIb supernova SN 1993J with some analysis of these spectra: the spectra include two HST spectra that have not been published before. The DIFF1 and DIFF2 formulae are presented in § 4, where we also present an example of fitted locally-normalized spectra. As an example of the utility of DIFF1/2, § 5 gives some preliminary statistical results for the spectra of supernovae of Types IIb, Ib, Ic, and Ic hypernovae. We will collectively refer to these supernovae as hydrogen-deficient core-collapse (HDCC) supernovae. Conclusions and discussion are given in § 6.

This paper makes use of and refers to material found at the first author’s online supernova spectrum database SUSPEND (SUpernovae Spectra PENDING further analysis). The URL is given in a footnote to the abstract.

2. LOCAL NORMALIZATION

The basic procedure of local normalization is to divide a spectrum by a smoothed version of itself where the smoothing length is sufficiently large that line features are largely erased, but the continuum shape is largely unaffected, in the smoothed version. The spectrum resulting from the division has ideally a continuum level of 1 everywhere as judged by the eye.

In § 2.1, we discuss the theoretical basis for local normalization. In § 2.2, we discuss our actual procedure for carrying out local normalization.
2.1. The Theoretical Basis for Local Normalization

For concreteness, in our discussion let us model an observed spectrum \( f_{\lambda, \text{obs}} \) by the following heuristic formula:

\[
f_{\lambda, \text{obs}} = S_{\lambda} f_{\lambda, \text{con}} f_{\lambda, \text{lin}} + C_{\lambda},
\]

where \( S_{\lambda} \) is some kind of multiplicative scaling error which could be the effect of unknown reddening or calibration error, \( f_{\lambda, \text{con}} \) is the intrinsic continuum spectrum, \( f_{\lambda, \text{lin}} \) is the line spectrum, \( C_{\lambda} \) is some extraneous and unknown contamination flux, and

\[
f_{\lambda, \text{int}} = f_{\lambda, \text{con}} f_{\lambda, \text{lin}},
\]

is the intrinsic spectrum. The line spectrum is defined by saying it is has a height of 1 when it is smoothed over some specified smoothing length that is of order the size of the full width of an individual line profile: thus

\[
\langle f_{\lambda, \text{lin}} \rangle = 1,
\]

where the angle brackets indicate smoothing. We make the assumption that \( S_{\lambda}, f_{\lambda, \text{con}}, \) and \( C_{\lambda} \) are all sufficiently slowly varying with wavelength that smoothing them over the the smoothing length leaves them effectively unchanged. (We are neglecting the possibility of light echo contamination (see § 1).)

Now say we smooth the observed spectrum over smoothing length. We obtain

\[
\langle f_{\lambda, \text{obs}} \rangle = \langle S_{\lambda} f_{\lambda, \text{con}} f_{\lambda, \text{lin}} \rangle + C_{\lambda},
\]

where we have used the assumption that \( C_{\lambda} \) is unaffected by smoothing. We assume no correlation between \( S_{\lambda} f_{\lambda, \text{con}} \) and \( f_{\lambda, \text{lin}} \). Thus

\[
\langle S_{\lambda} f_{\lambda, \text{con}} f_{\lambda, \text{lin}} \rangle = \langle S_{\lambda} f_{\lambda, \text{con}} \rangle \langle f_{\lambda, \text{lin}} \rangle = S_{\lambda} f_{\lambda, \text{con}},
\]

where we have used the assumptions that \( S_{\lambda} \) and \( f_{\lambda, \text{con}} \) are unaffected by smoothing. The locally-normalized spectrum \( f_{\lambda, \text{LN}} \) is given by

\[
f_{\lambda, \text{LN}} = \frac{f_{\lambda, \text{obs}}}{\langle f_{\lambda, \text{obs}} \rangle} = \frac{S_{\lambda} f_{\lambda, \text{con}} f_{\lambda, \text{lin}} + C_{\lambda}}{S_{\lambda} f_{\lambda, \text{con}} + C_{\lambda}} = f_{\lambda, \text{lin}} \left[ \frac{1 + C_{\lambda}/(S_{\lambda} f_{\lambda, \text{con}} f_{\lambda, \text{lin}})}{1 + C_{\lambda}/(S_{\lambda} f_{\lambda, \text{con}})} \right].
\]

What we would really like to have is the line spectrum (as we have defined it above) \( f_{\lambda, \text{lin}} \) and what we can obtain from measurements is \( f_{\lambda, \text{LN}} \). The \( f_{\lambda, \text{lin}} \) spectrum is what we believe contains much of the information about the object that we are interested in. The \( f_{\lambda, \text{lin}} \) and \( f_{\lambda, \text{LN}} \) spectra are equal if contamination is zero no matter what \( S_{\lambda} \) and \( f_{\lambda, \text{con}} \) may be given our assumptions. Thus, the effect of any scaling error is eliminated in absence
of contamination. This is one of the great benefits of using locally-normalized spectra in spectrum analysis.

If \( C_\lambda / (S_\lambda f_{\lambda,\text{con}}) \ll 1 \), then

\[
f_{\lambda,\text{LN}} \approx f_{\lambda,\text{lin}} \left[ 1 + \frac{C_\lambda}{S_\lambda f_{\lambda,\text{con}}} \left( \frac{1}{f_{\lambda,\text{lin}}} - 1 \right) \right],
\]

where we have expanded equation (6) to first order in \( C_\lambda / (S_\lambda f_{\lambda,\text{con}}) \). We see that contamination error only cancels to zeroth order. But if \( C_\lambda / (S_\lambda f_{\lambda,\text{con}}) \) is sufficiently small or line features are very weak (i.e., \( f_{\lambda,\text{lin}} \approx 1 \) in whichever of eq. (6) and (7) is relevant), then contamination will be a small error.

In the rest of this paper, we assume contamination is, in fact, negligible or has been adequately corrected for. Thus, we assume that the locally-normalized spectra are the line spectra as we have defined them above.

### 2.2. The Procedure of Carrying Out Local Normalization

There is probably no absolute optimum way of carrying out the basic procedure of local normalization. We have developed a prescription, however, that seems very adequate. The smoothing is box-car-like where we march through the spectrum wavelength by wavelength and make use of a box-car interval about each wavelength to construct a smoothed flux value at that wavelength. The size of the box-car interval is the smoothing length.

The choice of the smoothing length is a key point. If one makes the smoothing length too small, line features tend to get suppressed in the locally-normalized spectra: in the limit that the smoothing length goes to zero, the locally-normalized spectra become 1 everywhere. On the other hand, if one makes the smoothing length too large, then the locally-normalized spectra retain continuum features: in the limit that the smoothing length goes to infinity, the locally-normalized spectra are just ordinary (or globally) normalized spectra. Unfortunately in supernova spectra, the wavelength scale of the line features can be very large because of the large Doppler shift velocities that give rise to the line P Cygni profiles. (Supernova lines in the photospheric or optically-thick phase are broad P-Cygni lines with a blueshifted absorption and an emission feature centered roughly on the rest-frame line-center wavelength. See Jeffery & Branch (e.g. 1990, p. 173) for a discussion of P-Cygni line formation in supernovae.) Usually, the largest velocities one needs to consider are of order 30,000 km s\(^{-1}\) and these lead to relative Doppler shifts of line opacity of order \( v/c \sim 0.1 \) or 10%. However, line features can form at velocities of up to perhaps 50,000 km s\(^{-1}\) in Type Ic hypernovae (e.g., Mazzali et al. 2000) at early times and of up to at least 40,000 km s\(^{-1}\) in other kinds of supernovae if
seen at sufficiently early times in the UV as evidenced by the P Cygni line of the resonance multiplet Mg II $\lambda$2797.9 (e.g., Wiese et al. 1969, p. 30; Moore & Merrill 1968, p. 10) seen in the earliest IUE spectrum of Type II supernova SN 1987A (e.g., Pun et al. 1995). With Doppler shift velocity of 50,000 km s$^{-1}$, a relative wavelength shift would be $v/c \sim 1/6$ or 17%. Since one has to consider features shifted both to the blue and the red, relative the wavelength width of a line profile in extreme cases could be of order 34%. Continuum features that vary on the scale of 34% of wavelength certainly exist. Thus, for supernovae at least in extreme cases, one cannot completely decouple line and continuum behavior.

One must make a choice for the smoothing length for locally-normalized spectra that does not suppress some line behavior too much and that does not let too much continuum behavior leak in. There is probably no absolute optimum choice for all cases. We have found that a relative smoothing length of 30% of the current wavelength in the smoothing procedure is good: the smoothing region extends 15% blueward and 15% redward of the current wavelength as one marches through the wavelength points. The choice of 30% is motivated as follows. The continuum level determined by blackbody-behavior and reddening (Cardelli et al. 1989; O'Donnell 1994) is relatively constant over a region of 30% surrounding a wavelength. On the other hand, as mentioned above, usually the largest velocities one needs to consider are of order 30,000 km s$^{-1}$ and these lead to relative Doppler shifts of line opacity of order $v/c \sim 0.1$ or 10%: the whole line profile in this case will vary over 20% of the current wavelength. Thus a relative smoothing length of 30% should be effective at smoothing away most line features in the smoothed spectrum, but not smoothing the continuum shape too much.

To calculate the smoothed spectrum itself, a common approach is to simply numerically integrate (e.g., using the trapezoid rule) the spectrum over the box-car interval centered on the current wavelength and divide by the box-car interval to obtain an average flux for the box-car interval. For ordinary smoothing, nothing special needs to be done about the ends of the spectrum even though they can be treated somewhat wrongly. As an end is approached, the box-car interval gets progressively truncated at that end and the calculated average flux is progressively a less good approximation for the smoothed flux at the current wavelength: it is the smoothed flux for some point in the box-car interval farther from the end. In ordinary smoothing, the smoothing length is comparatively short and somewhat bad behavior at the ends of the spectrum is usually unimportant. But with a large smoothing length, such as we require for local normalization, this bad behavior at the ends can significantly degrade local normalization near the ends of the spectrum. If the spectrum continuum is relatively flat near the ends, there may be no noticeable degradation. But if the continuum is rising/falling at an end, then the smoothed spectrum becomes too small/large at that end and the locally-normalized spectrum continuum can appear to be rising/falling from 1 at the end. To avoid
this problem, we fit a line using least-squares to the spectrum in the box-car interval and determine the smoothed spectrum at the current wavelength from this fitted line. (Actually we fit a line to logarithmic flux as a function of logarithmic wavelength. This means we are fitting the flux to a power-law function. See below.) In the interior of the spectrum, this least-squares line fitting gives almost the same result as calculating an average flux for the box-car interval, but at the ends it prevents the locally-normalized continuum from departing obviously from 1. Since spectra are not necessarily very linear (or like a power-law function) over the box-car interval, it is not immediately obvious that this fitting of a line (actually a power-law function in our calculations) should always work well. But the results always seem pleasing to the eye and again we accept the eye’s judgment for our prescription.

In plotting spectra, we often prefer to plot logarithmic flux versus logarithmic wavelength. The choice of logarithmic wavelength owes to the fact that line feature widths are determined by the large Doppler shifts of opacity which are relative to the rest-frame line-center wavelength. For example, say a particular component of the line profiles forms in the ejecta moving at velocity $v$ along the line-of-sight counting velocity as positive in the direction of the observer. For given line of line-center wavelength $\lambda_{\text{line}}$, the component will appear in a spectrum at

$$\lambda = \lambda_{\text{line}}(1 - v/c),$$

assuming the 1st order Doppler formula (which is usually adequate for supernovae). Thus, for a set of lines, the components in a plot against wavelength will appear at different wavelength shifts from the line-center wavelengths. Thus, the line profiles will tend to have varying widths that scale with $\lambda_{\text{line}}$. In spectrum figures, this has an unpleasing effect of making the blue side of a spectrum look crowded and the red side look widely spaced. If one plots logarithmic wavelength, then for the aforementioned particular component one has a shift of

$$\log(\lambda) = \log[\lambda_{\text{line}}(1 - v/c)] = \log(\lambda_{\text{line}}) + \log(1 - v/c) \approx \log(\lambda_{\text{line}}) - \left(\frac{v}{c}\right)\log(e),$$

where we have expanded to 1st order in $v/c$ which is consistent with our assumption of the validity of 1st order Doppler shift formula. We see that in logarithmic wavelength, the shift of the particular component is the same for all line profiles. Thus, the line profiles will tend to have the same width across the spectrum. Varying line strength and line blending, of course, prevent the profiles from having an identical appearance. This argument for plotting logarithmic wavelength for ordinary spectra also applies to locally-normalized spectra as well.

In plotting ordinary spectra, we often make the choice of plotting logarithmic flux since it gives a more equal appearance to small and large features of a spectrum. Say we take
equation (2) (see § 2.1) for the intrinsic flux (which for the sake of argument we assume we know here) and take the logarithm: we get

$$\log (f_{\lambda,\text{int}}) = \log (f_{\lambda,\text{con}}) + \log (f_{\lambda,\text{lin}}).$$

(10)

From equation (10), we see that line profiles of the same relative size (i.e., the same $f_{\lambda,\text{lin}}$ profile) will have the same absolute size when logarithmic flux is plotted. Of course, varying relative line profile size and line blending will make the logarithmic line profiles vary. Nevertheless there is some equalization in line profile appearance in using logarithmic flux. Why do we prefer the more equal appearance of plotting logarithmic flux when large line features, of course, represent larger effects? The reason is that small line features present fine tests for spectrum modeling and clues about quantities (importantly composition) that may not be apparent in the large line features. Thus, giving more equal weighting to large and small line features makes sense.

Now in making a locally-normalized spectrum large and small line features tend to be equalized in size whether we start with ordinary flux or logarithmic flux, and so why use logarithmic flux in the local normalization procedure? We use logarithmic flux for creating locally-normalized spectra partially for consistency with our usual plotting practice for ordinary flux. However, we also plot the locally-normalized spectra on log-log plots, and using logarithmic flux in creating the locally-normalized spectra tends to make the locally-normalized continuum more like 1 everywhere to the eye on a log-log plot. The logarithmic flux and logarithmic wavelength of the spectra are actually used (as mentioned above) in making the smoothed spectrum when we fit a line to logarithmic flux as a function of logarithmic wavelength over the box-car interval (i.e., we fit flux with a power-law function over the box-car interval). Locally-normalized spectra calculated by fitting a line to flux as a function of wavelength (which we will call the unlogged locally-normalized spectra) look similar to those calculated with our logarithmic fitting, but the difference is clearly significant to the eye.

Before applying the large scale smoothing to effect the local normalization, we also apply to all our spectra a small scale box-car smoothing (using trapezoid rule integration) with a relative smoothing length of $\delta \lambda / \lambda = 1/300$ which corresponds to a Doppler shift velocity of about $1000 \text{ km s}^{-1}$. This smoothing reduces the noise in some of the spectra particularly at the spectrum ends. For some particularly noisy spectra, we sometimes use a larger smoothing length. The small scale smoothing should not degrade the spectrum features very much. Supernova line features generally vary relatively slowly over wavelength shifts corresponding to velocity changes of order $1000 \text{ km s}^{-1}$. Interstellar and telluric lines can vary on shorter wavelength scales, but these are not intrinsic to supernova spectra and could be corrected for if necessary.
A stand-alone fortran-95 code locnorm.f for making locally normalizing spectra is available for downloading from SUSPEND under the heading Example Programs. All the spectra currently in SUSPEND are shown plotted in original and locally-normalized format under the heading Supernovae by Epoch in html files in supernova directories.

3. EXAMPLES OF LOCALLY-NORMALIZED SUPERNOVA SPECTRA

Figures 1, 2, 3, and 4 (which we discuss individually below in separate subsections) show locally-normalized supernova spectra plotted on log-log plots together with the original spectra in the $f_{\lambda}$ representation (i.e., the flux per unit wavelength representation): the locally-normalized spectra obviously have a continuum level of about 1 and the original spectrum have been vertically shifted to be well displayed. The spectra are all for Type IIb supernova SN 1993J which occurred in M81 (NGC 3031). They have been deredshifted using host galaxy heliocentric velocity $-39 \pm 2$ km s$^{-1}$ from Leda (Paturel et al. 2003). The original spectra have not been dereddened. The $E(B - V)$ reddening value for SN 1993J is quite uncertain. Richmond et al. (1994) after a lengthy consideration of many methods of determining $E(B - V)$ suggest that it is in the range $\sim 0.08$–$0.32$ mag which is a range from low to moderate reddening. Thus, the intrinsic color $B - V$ for any epoch has value range of about $\sim 0.24$ mag.

Because M81 is at a distance of only $3.63 \pm 0.34$ Mpc (as determined from Cepheids (Freedman et al. 1994)), SN 1993J became a very bright and well-observed supernova. It was discovered on 1993 March 28.91 UT (JD 2449075.40) (Ripero 1993) which was probably only about a day after explosion (or core collapse) which may have been 1993 March 27.9 ± 0.2 (JD 2449074.4 ± 0.2) (Richmond et al. 1994; Clocchiatti et al. 1995). An initial very high peak in optical brightness was partially observed (e.g., Richmond et al. 1994): this peak, which probably happens for all core-collapse supernovae soon after shock break-out, is usually unobserved, and so is not used to define conventional maximum light even though it may often/always be higher than conventional maximum light as it is for SN 1993J. SN 1993J’s (conventional) maximum light occurred on 1993 April 16 (JD 2449093.7) in $B$, 1993 April 17 (JD 2449095.0) in $V$, and 1993 April 17 ± 1 (JD 2449094.7 ± 1.0) in UVOIR bolometric luminosity (Richmond et al. 1994). The UVOIR bolometric luminosity rise time to maximum light was about 20 days (Richmond et al. 1994). The maximum $B$ and $V$ apparent magnitudes were, respectively, $11.35 \pm 0.05$ and $10.80 \pm 0.05$ (Richmond et al. 1994).

SN 1993J was originally classified as a Type II supernova because it had conspicuous hydrogen Balmer lines. However, these lines did not become as strong as in typical Type II
supernovae and by about 40 days after explosion SN 1993J evolved to resemble Type Ib supernovae (e.g., Richmond et al. 1994). Type Ib supernovae have metal lines and conspicuous helium lines and, it seems in general, hydrogen lines that are somewhat weak and hard to identify especially after maximum light (Branch et al. 2002). Supernovae that undergo a transformation from Type II to Type Ib have come to be called Type IIb’s using a term invented by Woosley et al. (1987) in a theoretical prediction that small-hydrogen-envelope supernovae could exist. Type IIb’s are considered to be hydrogen-deficient compared to ordinary Type II supernovae, but have more hydrogen than Type Ib/c’s. The observational prototype Type IIb is SN 1987K (Filippenko 1988). Two other Type IIb supernovae with a significant number of published spectra are SN 1996cb (Qiu et al. 1999; Matheson et al. 2001) and SN 1998fa (Matheson et al. 2001). As of 2006 July 4, the supernova list of Central Bureau for Astronomical Telegrams (2006) contains 21 Type IIb or possible Type IIb supernovae (including SN 1987K which is actually listed as a Type II).

The ejecta of SN 1993J had a shock interaction with a thick circumstellar wind shed by the supernova progenitor as evidenced by radio and X-ray emission (e.g., Fransson et al. 1996, and references therein). In this respect, SN 1993J was like the Type II supernovae SN 1979C and SN 1980K which also showed strong radio emission (e.g., Weiler et al. 1986). The interaction also had an effect on the UV spectra of SN 1993J similar to that on SN 1979C and SN 1980K. We will discuss this briefly in § 3.2.

We have indicated some line identifications in the figures. For clear P Cygni lines, we put the identification labels with the blueshifted absorption features of the P Cygni lines since this is usually the most distinct part of the line profiles. For emission lines, we put the identification labels with the emission features of the emission lines. P Cygni lines dominate optical/IR spectra in the photospheric phase and emission lines in the nebular phase. Figures 1 and 2 show photospheric phase spectra and Figures 3 and 4 show nebular phase spectra.

We should remark that in the photospheric phase, a supernova is optically thick. As the ejecta expands, the opacities must fall with density and the supernova gradually enters the nebular phase where the ejecta is optically thin in the optical/IR although it remains optically thick in the UV for much longer where there are very strong metal lines particular from the iron-peak elements. There is no sharp dividing time between the two phases. One often just starts using the term nebular phase when the emission features start to dominate the optical/IR line pattern. There is also no sharp distinction between P Cygni lines (possessing a blueshifted absorption and a rest-frame line-center emission) and emission lines (which just have rest-frame line-center emission). In practice, when the absorption feature of a P Cygni line has become hard to identify, then we can say it has become an emission line.
We should also remark that supernova matter after the explosion phase is in uniform motion and the whole ejecta structure just scales up linearly with time $t$ since the explosion which is effectively a point event. This motion is called homologous expansion. In homologous expansion, the radial velocity of any matter element of ejecta is a good comoving-frame radial coordinate and we use it as such as is customary in supernova studies. The densities of the matter elements in the ejecta scale as $t^{-3}$.

3.1. Figure 1: the 1993 March 31 Spectrum

Because of its closeness and brightness, SN 1993J was spectroscopically very well observed with the earliest spectrum taken on 1993 March 29.88 UT (JD 2449076.38) with the Isaac Newton Telescope by E. Perez and D. Jones (Clocchiatti et al. 1995). The spectrum in Figure 1 is also an early one from 1993 March 31 (JD 2449078.35) (Barbon et al. 1995) which is about 4 days after explosion and is about 16 days before UVOIR bolometric maximum light. The original spectrum rises very steeply to the blue showing that the supernova photosphere was very hot at this phase. From Figure 11 of Richmond et al. (1994), we estimate that blackbody fit to the photometry for this phase would give a spectral temperature in the range $\sim 11000$–$18000$ K where the range is caused by the uncertainty in the reddening. The uncertainty in spectral temperature highlights the difficulty in relying on continuum shape for determining the physical properties of supernovae from heterogeneous types with uncertain reddening and supports our argument for giving more weight to the line pattern.

We have identified the lines in Figure 1 mostly following Baron et al. (1995). For this spectrum and the others we also rely on our general understanding of spectrum formation in making identifications. The H$\delta$ identification is perhaps a bit uncertain because it seems rather weak. The He I line is He I $\lambda$5875.7 (e.g., Wiese et al. 1966, p. 14) and the Ca II line actually arises from the Ca II IR triplet Ca II $\lambda\lambda$8579.1 (where we have cited the $gf$-weighted mean line wavelength: the component lines are at 8498.02 Å, 8542.09 Å, and 8662.14 Å) (e.g., Wiese et al. 1969, p. 251). (Hereafter we usually cite only the $gf$-weighted mean line wavelength for multiplets.) The line velocities (i.e., the velocities corresponding to the Doppler blueshifts of the line absorption minima from the rest-frame line-center wavelengths) are about 11000 km s$^{-1}$ (H$\delta$), 5900 km s$^{-1}$ (H$\gamma$), 13400 km s$^{-1}$ (H$\beta$), 12800 km s$^{-1}$ (He I), 12700 km s$^{-1}$ (H$\alpha$), and 26800 km s$^{-1}$ (Ca II). We use the $gf$-weighted mean wavelengths for calculating the line velocities of multiplets. (Here and throughout our discussion wavelength, measurements are made from the locally-normalized spectra and we do a little smoothing by eye to determine the minima when needed.) The low velocity of H$\gamma$ may be because of line blending with H$\delta$ since the lines are separated in Doppler shift velocity by only 17000 km s$^{-1}$. 
The much higher velocity for the Ca II line than for the others makes the identification a bit uncertain, but there seems little else it could be.

A remarkable feature of the line pattern is that the H\(\beta\) line is much stronger than the H\(\alpha\) line: this cannot happen in local-thermodynamic equilibrium (LTE). The feature may require a non-LTE (NLTE) explanation. On the other hand, the feature may be a result of an observational error of some kind since a spectrum from the same day reported by Baron et al. (1995) does not clearly show the H\(\beta\) line stronger than the H\(\alpha\) line. In any case, Baron et al. (1995) using the NLTE code PHOENIX were able to fit the continuum well, but not the lines for this phase of the supernova. Their difficulty may lie in the uncertain physical structure of their model and/or the uncertain reddening.

3.2. Figure 2: the 1993 April 15 Spectrum

Figure 2 shows the SN 1993J spectrum from 1993 April 15 (Jeffery et al. 1994): this was 2 days before UVOIR bolometric maximum light and about 18 days after explosion. The spectrum is a combination of a Hubble Space Telescope (HST) UV-blue-optical spectrum and Lick Observatory ground-based spectrum. The HST spectrum was obtained as part of the HST Supernova INtensive Study (SINS) General Observer program (Kirshner et al. 1988). The two spectra agreed well in their overlap region 3120–3276 Å, and so we simply joined them at 3240 Å cutting off the overlapping ends. The line identifications mostly follow from Jeffery et al. (1994) and Baron et al. (1995). The line velocities are in the range \(\sim 7000–11000 \text{ km s}^{-1}\) (Jeffery et al. 1994). We emphasize though that the lines are often blended and the identifications in many cases only indicate the strongest contributor to the line feature.

The 4 strongest hydrogen Balmer lines, the He I \(\lambda 5875\) line and the Ca II \(\lambda 8579\) line are again present. We note that H\(\alpha\) is now the strongest line in the optical part of the spectrum and that it is in net emission which is typical of H\(\alpha\) lines in Type II supernovae.

The absorption centered at about 3835 Å is caused by the Ca II H&K lines: i.e., the multiplet Ca II \(\lambda 3945.2\) (citing the \(gf\)-weighted mean wavelength (e.g., Wiese et al. 1969, p. 252)). This multiplet is very strong because it is a resonance multiplet (i.e., it arises from the ground level). The line velocity for the Ca II H&K lines is 8360 km s\(^{-1}\). The Fe II lines mainly arise from the multiplets Fe II \(\lambda 4555\), Fe II \(\lambda 4561\), and Fe II \(\lambda 5060\) (citing the \(gf\)-weighted mean wavelengths (e.g., Kirshner et al. 1993, Table 4)). The absorption centered at about 7570 Å may be the absorption of the line caused by the multiplet O I \(\lambda 7773.4\) (e.g., Wiese et al. 1966, p. 79). This line is common in Type Ib/c supernovae (e.g., Branch et al. 2002) and Type Ia supernovae (e.g., Kirshner et al. 1993), but unfortunately its absorption usually coincides with a strong telluric line and it is sometimes, as here, unclear if the
telluric line has been adequately corrected for. We regard this identification as tentative. If the identification is correct, the O I λ7773.4 line velocity is 7830 km s$^{-1}$. By about 40 days past UVOIR bolometric maximum light, the O I λ7773.4 line seems clearly present (Barbon et al. 1995).

The UV of SN 1993J is similar to that of SN 1979C and SN 1980K and is probably greatly modified from what the bare supernova would yield by the circumstellar shock interaction (Jeffery et al. 1994). In a series of papers, Fransson (1982, 1984a,b) has given a detailed model of supernova UV behavior with circumstellar interaction: this work was motivated by International Ultraviolet Explorer ($IUE$) data for SN 1979C and SN 1980K. In brief, some fraction of UVOIR light from the supernova photosphere is scattered by shock-heated electrons in the shock region of ejecta and circumstellar matter and Comptonized (i.e., blueshifted up to UV/extreme UV). Some of this Comptonized light just escapes the supernova environment, but some back-heats the outer supernova ejecta and creates a complex, layered, photoionization region above the supernova photosphere. The emission from this ejecta photoionizes the direct emission from the supernova interior that emerges at the photosphere. In the optical, this extra-emission is probably negligible (except in some lines), but in UV blueward of perhaps $\sim 2800\,\AA$ it appears to become increasingly important by comparison to spectra from supernovae with little or no circumstellar interaction (Jeffery et al. 1994). As well as continuum emission from recombination and free-free emission, emission lines from ionized atoms can make a considerable contribution: the emission lines are primarily resonance lines that have been collisionally excited by the Compton-flux heated medium. The layers of Comptonized-flux photonization are geometrically quite narrow, and thus the observer sees emission primarily from a geometrically thin spherical shell of ejecta moving at velocities of order $10^4\,\text{km s}^{-1}$. Below the shell is the optically thick photosphere. Thus, the observer sees mostly emission from the near-side of the ejecta with line-of-sight velocities ranging from much less than $10^4\,\text{km s}^{-1}$ away from the observer up to line-of-sight velocities of order $10^4\,\text{km s}^{-1}$ toward the observer. For emission lines, the varying Doppler shift of the moving shell will result in broad blueshifted emission features. The Comptonized flux dominates the line emission for of order 30 days after explosion and then X-ray emission from the circumstellar interaction begins to dominate.

Another effect which enhances the blueshift of line emission from a shell is diffuse reflection from the photosphere of line emission: this diffusely reflected light can only come toward the observer from the near-side of the photosphere (Chugai 1988). The photons that reflect off the core in the direction of the observer are given an extra blueshift in the observer-frame by the line-of-sight motion of the photosphere matter. The reflection can be caused by scattering from electrons, but in the UV, particularly in the nebular phase when the electron optical depth through the ejecta has become small, the reflection is probably mainly by
scattering from the many resonance lines or lines arising from low-energy metastable levels of iron peak elements.

We have not attempted a full identification of features seen in the UV region of the spectra we present. However, the strong emission-like feature peaking at 2717 Å in the spectrum must be the resonance multiplet Mg II λ2797.9 (e.g., Wiese et al. 1969, p. 30; Moore & Merrill 1968, p. 10) blueshifted and in emission and produced as discussed by Fransson. (The emission features analyzed by Fransson were primarily those from blueward of 2000 Å, but his analysis seems to apply to Mg II λ2797.9 as well.) A similar Mg II λ2797.9 emission appeared in SN 1979C about 7 days after optical maximum light (Panagia et al. 1980) and was still present in the last IUE observation about 101 days after maximum light (INES 2006). (Optical maximum light for SN 1979C was about 1979 April 15 (Barbon et al. 1982).) In SN 1980K, a similar Mg II λ2797.9 emission seems to have arisen about 26 days after optical maximum light and be present on the last IUE observation from about 62 days past maximum. The SN 1980K spectra seem to have poorer quality and the emission is generally less clear than in SN 1979C, but it seems definite about 35 days after maximum (INES 2006). (Optical maximum light of SN 1980K was about 1980 November 4 (Buta 1982).) We note that an Mg II λ2797.9 emission line is a common feature of quasar emission-line clouds (e.g., Reichard et al. 2003) which somewhat resemble the UV emission shell of circumstellar-interacting supernovae (Fransson 1984b).

The Mg II λ2797.9 emission peak wavelength 2717 Å corresponds to a Doppler blueshift velocity of 8700 km s\(^{-1}\) which is within the range of P-Cygni absorption trough velocities in the optical \(\sim 7000–11000\) km s\(^{-1}\) (Jeffery et al. 1994). If we just define 2700 Å as the characteristic blue edge of the Mg II λ2797.9 emission (which is good for later phases: see §§ 3.3 and 3.4 below), the corresponding Doppler blueshift velocity is 10500 km s\(^{-1}\). We take this velocity as the characteristic velocity of the Mg II λ2797.9 emission shell.

Without strong circumstellar interaction, the Mg II λ2797.9 multiplet would probably contribute along with Fe II lines to a blueshifted P-Cygni absorption feature. For Type Ia supernovae, the absorption can be seen in the HST spectra of SN 1992A in the photospheric phase (Kirshner et al. 1993) and, perhaps, well into the nebular phase (291 days after \(B\) maximum light) although not so identified (Ruiz-Lapuente et al. 1995). For Type II supernovae, the absorption can be seen in the IUE spectra of SN 1987A (Pun et al. 1995). The difference between having emission and absorption is probably explained by saying that with circumstellar interaction the line-forming region is hotter than the inner ejecta and without circumstellar interaction it is colder.

No other identifications in the UV of SN 1993J are probably possible without a more detailed analysis.
3.3. Figure 3: the 1993 September 15.5 Spectrum

Figure 3 shows the SN 1993J spectrum that we have constructed from an 1993 September 17 \textit{HST} UV-blue-optical spectrum and a 1993 September 14 ground-based spectrum. We will call the combined spectrum the 1993 September 15.5 spectrum: thus it comes from about 151 days after UVOIR bolometric maximum light and 171 days after explosion. Both component spectra were obtained as part of the SINS program: the \textit{HST} spectrum has not been published before. The \textit{HST} spectrum has not been given a definitive flux calibration (Challis 1994) and we are unsure about the ground-based spectrum. The two spectra agreed well in shape in the region 4250–4781 Å, and so we simply joined them at 4253.6 Å cutting off the overlapping ends. The ground-based spectrum had to be scaled by a factor of 0.35.

The supernova at the time of this spectrum is clearly in the nebular phase because of the strong optical/near-IR emission lines. The line identifications for the 3 strongest emission lines in the optical/near-IR are based on analysis of Houck & Fransson (1996) for a 1993 August 15 SN 1993J spectrum: this extrapolation is reliable because these lines are strong and well identified in other HDCC supernovae. The actual lines are the forbidden multiplet \([\text{O I}] \lambda\lambda 6300, 6364\) which is a resonance multiplet (e.g., Wiese et al. 1966, p. 82) and the forbidden multiplet \([\text{Ca II}] \lambda\lambda 7291, 7324\) (e.g., Wiese et al. 1969, p. 255) and, again, the Ca II IR triplet.

The absorption with minimum at 3829 Å is caused by Ca II H&K lines (multiplet Ca II \(\lambda 3945.2\)). The line velocity of the Ca II H&K lines is 8800 km s\(^{-1}\). There is no obvious strong emission feature associated with these lines. Evidently, the Ca II H&K lines are optically thick and strongly trap photons in their resonance regions. The trapping process is as follows. Photons redshift into resonance with the lines and are absorbed and then resonantly emitted, but cannot easily redshift out of resonance because of continually being re-absorbed after redshifting very little. So there are many re-absorptions and re-emissions. Much of the photon energy ultimately escapes by being emitted by the Ca II IR triplet which shares the same upper level with the Ca II H&K lines (e.g., Moore & Merrill 1968, p. 12): this energy leak mostly suppresses the emission feature of the Ca II H&K lines. The lower level of the Ca II IR triplet is metastable and the upper level of the forbidden multiplet \([\text{Ca II}] \lambda\lambda 7291, 7324\) (e.g., Moore & Merrill 1968, p. 12). Collisions are weak in the nebular phase due to low density, and hence the decay from the metastable level becomes increasingly radiative which results in the strong \([\text{Ca II}] \lambda\lambda 7291, 7324\) emission feature. In brief, Ca II is showing strong fluorescence.

In the UV, the blueshifted Mg II \(\lambda 2797.9\) emission line now peaks at 2726 Å which corresponds to Doppler blueshift velocity of 7700 km s\(^{-1}\) and has now increased in height to about 3.4 times the continuum level. It overall shape, however, is much the same as earlier
(see Figure 5 in § 3.4).

The peak centered at 4563 Å may be at least in part the emission feature of the semi-forbidden line Mg I$\lambda4571$ (e.g., Moore & Merrill 1968, p. 16; Wiese et al. 1969, p. 26; NIST 2006). The emission is relatively weak, but Houck & Fransson (1996) predict Mg I$\lambda4571$ as a weak emission line (seemingly blended with some other emission line) for 1993 August 15 and believe the Mg I$\lambda4571$ emission should be become stronger with time. If the 4563 Å emission is Mg I$\lambda4571$, then the line has been Doppler blueshifted by velocity 520 km s$^{-1}$. The 1993 November 15.5 spectrum in Figure 4 (see § 3.4) shows stronger emission peaking at 4551 Å and if this is caused by Mg I$\lambda4571$, then it has been Doppler blueshifted by velocity 1050 km s$^{-1}$. We note that the [O I]$\lambda6300$ emission peaks at 6302 Å (corresponding to a Doppler redshift velocity of 95 km s$^{-1}$) in the 1993 September 15.5 spectrum and at 6300 Å (corresponding no Doppler shift at all) in the 1993 November 15.5 spectrum. In order for the 4563 Å and 4551 Å emissions to be primarily Mg I$\lambda4571$, there must be some blueshifting asymmetry or significant continuum optical depth. The lack of significant shifts in the [O I]$\lambda6300$ emission peaks argues for symmetric emission and optically thin ejecta in the optical. There is the complication, however, that the blue side of the optical generally has more strong lines than the red side of the optical and may persist being somewhat optically thick longer than the red side. The upshot of all these considerations is that we draw no conclusion about the identities of the 4563 Å and 4551 Å emissions. These emissions may be blueshifted Mg I$\lambda4571$ or Mg I$\lambda4571$ blended with something else, but more evidence is needed.

(The case for Mg I$\lambda4571$ identification in the optical is different than for the Mg II$\lambda2797.9$ identification in the UV. We expect the nebular-epoch ejecta to be optically thin in the optical and optically thick in the UV at least for a long time. The UV optical thickness in the nebular epoch is the cause of the blueshift of the Mg II$\lambda2797.9$ emission as we argued in § 3.2.)

We note that the strong emission features in the optical/near-IR in both versions of the spectrum, but particularly in the locally-normalized version, seem to be adjacent to absorptions. There may be real strong absorptions at these locations from lines that are optically thick, but a synthetic spectrum analysis is probably needed to show that. The locally-normalized spectra for the nebular phase, however, may be a bit deceptive. The process of local normalization will tend to make emission features sit in troughs that are below 1, and thus create the appearance of adjacent absorptions. In reality, the emission lines may just be added to a fairly smooth continuum that holds across the spectrum (or at least the optical/IR spectrum) and there may be no significant absorptions (at least in the optical/IR part). So in regard to nebular spectra, local normalization as we have implemented it may not be ideal in that it probably cannot scale a real physical nebular
continuum to 1 at least in the presence of strong emission features. On the other hand, there may be no well-defined physical continuum level in which case applying local normalization does not obviously worsen the appearance of the spectra for analysis. Our arguments for local normalization based on uncertain reddening and calibration are still valid for the nebular phase.

For the photospheric phase, local normalization is probably much better than for the nebular phase. First, there is a physical continuum in the spectrum (although it almost never a pure blackbody continuum). Second, the smoothing of absorptions and emissions to create the smoothed spectrum is an averaging that must draw the smoothed spectrum close to the physical continuum level. The locally-normalized spectrum will then have a physical continuum close to 1.

3.4. Figure 4: the 1993 November 15.5 Spectrum

Figure 4 shows the SN 1993J spectrum that we have constructed from 1993 November 14 HST UV-blue-optical spectrum and 1993 November 17 Lick Observatory ground-based spectra obtained as part of the SINS program: the HST spectrum has not been published before; the Lick spectrum was taken by A.V. Filippenko and T. Matheson (Filippenko 1994). We will call the combined spectrum the 1993 November 15.5 spectrum: thus it comes from about 212 days after UVOIR bolometric maximum light and 232 days after explosion. The HST spectrum has not been given a definitive flux calibration (Challis 1994) and the ground-based spectrum is not absolutely calibrated, although its relative calibration is excellent (Filippenko 1994). The two spectra agree well in shape in the overlap region 3120–4781 Å, and so we simply joined them at 4200 Å cutting off the overlapping ends. The ground-based spectrum had to be scaled by a factor of $5.55 \times 10^{-9}$ after converting it from the frequency representation to account for unit conversion. The HST spectrum blueward of $\sim 2425$ Å seems to decline in a manner too steep to be physically real and we do not trust it (nor the locally-normalized spectrum we derive from it) there.

Qualitatively, the 1993 November 15.5 spectrum is much like the 1993 September 15.5 spectrum. (We are relying on the locally-normalized spectra for comparisons.) The Mg II $\lambda 2797.9$, [O I] $\lambda\lambda 6300,6364$, and [Ca II] $\lambda\lambda 7291,7324$ emissions have increased in height relative to the continuum and the Ca II IR triplet emission has decreased. The emission at 4563 Å (whatever it may be) in the 1993 September 15.5 spectrum has increased and shifted to 4551 Å. The Ca II H&K absorption is not as deep. The peak of the Mg II $\lambda 2797.9$ emission is at 2726 Å just as in the 1993 September 15.5 spectrum.
In Figure 5, we show the time development of Mg II $\lambda 2797.9$ emission in locally-normalized spectra. Although the emission line scales up with time, its appearance is remarkably stable. Our definition of 2700 Å as the characteristic blue edge of the Mg II $\lambda 2797.9$ emission (with corresponding Doppler blueshift velocity $10500 \text{ km s}^{-1}$) is valid for all the phases we display. The behavior of the spectra in the region $\sim 2780$–2820 Å is not very certain since we have just omitted wavelength points that correspond to the interstellar Mg II $\lambda 2797.9$ lines which arise from gas clouds along the line of sight to the supernova. Such clouds, identified from interstellar lines of various ions, have heliocentric velocities in the range from about $-140$ to $181 \text{ km s}^{-1}$ (Wamsteker et al. 1993; Wheeler et al. 1993).

Later HST UV spectra of SN 1993J from day 649 to about day 2560 after UVOIR bolometric maximum light show the Mg II $\lambda 2797.9$ emission becoming more box-like in appearance and more symmetrical about the rest-frame line-center wavelength with a Doppler velocity half-width of $\sim 10000 \text{ km s}^{-1}$ (Fransson et al. 2005). It seems that the circumstellar interaction back-heating persists for a long time, but that the optical depth through the ejecta diminishes and thus occultation of the far side of the Mg II $\lambda 2797.9$ shell decreases. The flattish top of the late emission line is characteristic of emission lines from expanding, optically thin shells as first shown by Beals (1931); other references for the flat-top emission from expanding shells are Mihalas (1978, p. 477) and Jeffery & Branch (1990, p. 190).

3.5. The Reliability of Local Normalization with Continuum Distortion

The main goal of local normalization is to eliminate variations in the continuum in comparing spectra. It is very relevant then to ask if a spectrum with its continuum distorted in various ways is always reduced to the same locally-normalized spectrum: i.e., to ask if local normalization is reliable. To test if this is the case, we have weighted each of the four original spectra in this section by $\lambda^2$ (effectively converting the spectra to the $f_\nu$ representation) and $\lambda^4$ and then locally normalized these weighted spectra.

For each original spectrum, a plot (using an expanded scale compared to Fig. 1–4) of the locally-normalized spectra calculated with and without weights shows exact overlap to the eye of the locally-normalized spectra. On the other hand, unlogged locally-normalized spectra (see § 2.2) were distinct when plotted: they generally ran lower than the other locally-normalized spectra.

As mentioned in the Introduction (§ 1), DIFF1 is essentially a mean relative difference between the line patterns of two spectra and thus the reader can roughly understand DIFF1 values before seeing the precise formula (which is given below in § 4.1). The DIFF1 values
for spectrum pairs consisting of the locally-normalized weighted spectra and the locally-normalized unweighted spectra were of order 0.001–0.0015 for weight $\lambda^2$ and of order 0.002–0.003 for weight $\lambda^4$. These are very small mean relative differences. On the other hand, tests of the unlogged locally-normalized spectra with the other locally-normalized spectra gave DIFF1 values of order 0.4–0.7 which are comparable to DIFF1 values between spectra of the same supernova from phases differing by several days.

We conclude that local normalization is reliable in reducing continuum-distorted versions of the same spectrum to the same locally-normalized spectrum. But different local normalization procedures will produce noticeably distinct locally-normalized spectra. Thus, one should choose a single local normalization procedure and stick with it.

4. DIFF1 AND DIFF2

In this section, we introduce the formulae for DIFF1 (§ 4.1) and DIFF2 (§ 4.2), show an example of two spectra fitted by minimizing the DIFF2 value (§ 4.3), and compare DIFF1 and DIFF2 (§ 4.4).

4.1. DIFF1

Say $f_i$ is the locally-normalized flux at wavelength $\lambda_i$. Let

$$\delta_i = f_i - 1$$

be the flux difference at wavelength $\lambda_i$. The flux difference is the relative difference of the line pattern from the continuum or absolute difference of the line pattern in a locally-normalized spectrum. We define DIFF1 between a spectrum 1 and a spectrum 2 by the formula

$$\text{DIFF1} = \frac{1}{I} \sum_i \frac{|\delta_{1,i} - \delta_{2,i}|}{\max(|\delta_{1,i}|,|\delta_{2,i}|,\xi)},$$

where the subscripts 1 and 2 indicate the spectrum and $I$ is the total number of wavelength points in the summation. The $\xi$ is a small number that we set to $10^{-15}$. This is typically the smallest number that added to 1 that creates a number significantly different from 1 with fortran 95 double precision arithmetic. The $\xi$ prevents arithmetic exceptions, but is only very rarely invoked.

From equation (12), it is clear why we describe DIFF1 as the mean relative difference in wavelength pattern between two spectra (see §§ 1 and 3.5). If DIFF1 $<< 1$, the spectra
will be much alike. If $\text{DIFF1} \gtrsim 1$, the spectra will be substantially different. In fact, the DIFF1 function has an upper limit:

$$|\delta_{1,i} - \delta_{2,i}| \leq |\delta_{1,i}| + |\delta_{2,i}| \leq 2 \max (|\delta_{1,i}|, |\delta_{2,i}|) ,$$

and thus

$$\text{DIFF1} \leq 2 .$$

We would not usually expect a DIFF1 value approaching 2 since that would mean that two input spectra had line patterns that were nearly mirror images of each other about the continuum level. In fact, any DIFF1 values substantially over 1 would show an anticorrelation between spectrum line patterns. There is no strong physical reason to expect much anticorrelation of spectrum line patterns, but some random, accidental anticorrelation should happen, and so some DIFF1 values over 1 will occur when applying DIFF1 to heterogeneous samples of spectra.

We have subjected the set of all possible spectrum pairs drawn from the sample of HDCC supernovae described in § 5 to the DIFF1 test. (There are 16 supernovae, 156 spectra, and 23724 valid spectrum pairs: some possible pairs are excluded as discussed below.) The resulting distribution of DIFF1 values has mean, standard deviation, minimum value, and maximum value of, respectively, 0.868, 0.139, 0.244, and 1.320 (where we have reported more digits than are significant to allow for numerical consistency checks). That the mean 0.868 is so close to 1 shows that spectrum pairs on average are not much alike. The minimum 0.244 suggests that even spectra from the same supernova at close to the same phase vary from each other significantly. (We have excluded redundant spectra from the same phase period (which we usually set to being 1 day) for a supernova (as described in § 5) since those should be nearly identical to the included spectrum aside from observational error.) As we expected, the maximum DIFF1 value is much less than 2.

We note that if we used $f_i$’s in the formula for DIFF1 (i.e., eq. (12)) instead of $\delta_i$’s (which would only change denominators, in fact), we would have lessened the sensitivity to line patterns. For example, Type IIb and Type Ic hypernovae at early phases (i.e., well before maximum light) both have relatively weak lines in the optical. In the case of Type IIb supernovae, this is because they have hydrogen-dominated atmospheres that when hot (i.e., over of order 10000 K) show weak lines. (Why the lines should be as weak as they are in early Type IIb spectra has not been fully theoretically elucidated in NLTE (Baron et al. 1995).) Type Ic hypernovae show weak lines at early times because they are hot and their velocity structure is very fast giving them extremely broad P-Cygni lines with line velocities of up to 30000 km s$^{-1}$ and perhaps higher (e.g., Mazzali et al. 2000; Patat et al. 2001). The stretching out of P-Cygni lines by very high velocities makes them shallower and more overlapping. Both effects lessen the scale size of the lines relative to the continuum. If we used $f_i$’s
in the definition of DIFF1 instead of $\delta_i$'s, the DIFF1 values between early phase spectra of Type IIb's and Type Ic hypernovae would be small even though the spectra and the supernovae are intrinsically quite different because the numerators would be relatively small and the denominators relatively large. However, the given formula for DIFF1 distinguishes the two types of spectra because the denominators become small when the lines features are small.

For our calculations, we interpolated the locally-normalized spectra onto a grid of equally spaced points in logarithmic wavelength. For coding convenience, we chose $\log(1 \text{ Å}) = 0$ to be the zero point of the logarithmic wavelength scale. For the grid increment, we chose

$$\Delta \log(\lambda) = \log(e) \times \Delta \ln(\lambda) = \frac{\log(e)}{3000},$$

(15)

where $\Delta \ln(\lambda) = 1/3000$. Since

$$\Delta \ln(\lambda) = \ln(\lambda_i) - \ln(\lambda_{i-1}) = \ln \left( \frac{\lambda_i}{\lambda_{i-1}} \right) = \ln \left( 1 + \frac{\Delta \lambda_i}{\lambda_{i-1}} \right) \approx \frac{\Delta \lambda_i}{\lambda_{i-1}},$$

(16)

for $\Delta \ln(\lambda) \ll 1$, the logarithmic wavelength increment corresponds to relative change in wavelength for each increment of approximately $1/3000$: thus, near 3000 Å, the increment is about 1 Å; near 6000 Å, about 2 Å. The Doppler shift velocity corresponding to a relative wavelength increment of $1/3000$ is about $100 \text{ km s}^{-1}$. As mentioned in § 2.2, supernova line features generally change relatively slowly over wavelength shifts corresponding to velocity changes of order $1000 \text{ km s}^{-1}$, and thus our grid should be sufficiently fine for almost all supernova spectra. As also mentioned in § 2.2, interstellar and telluric lines can vary on shorter wavelength scales, but these are not intrinsic to supernova spectra and could be corrected for if necessary.

The range of wavelengths we include in the DIFF1 test includes all of the wavelength overlap region of the two spectra where the two spectra seem physically good: i.e., we exclude spectrum ends that are too noisy or appear to be unphysical through some calibration error which is fairly common at the ends of observed spectra. We only consider DIFF1 tests valid if the logarithmic wavelength range overlap is $\geq 0.2 \times \log(e)$ which corresponds to a relative wavelength overlap of about 20%. Since almost all the spectra we consider cover at least the range 4000–7000 Å, almost all DIFF1 evaluations are considered valid.
4.2. DIFF2

The DIFF2 formula is almost the same as equation (12) for the DIFF1 formula:

\[
\text{DIFF2} = \frac{1}{I} \sum_i \frac{|\delta_{1,i} - \delta_{2,i+k}|}{\max(\{|\delta_{1,i}|, |\delta_{2,i+k}|, \xi\})}.
\]

The only difference is an additive term \(k\) in the subscript which is chosen to minimize the DIFF2 value. We allow up and down logarithmic wavelength shifts of only up to \(0.05 \times \log(e)\). Thus, we allow up and down relative wavelength shifts of only about 5%. Given that \(\Delta \lambda_i/\lambda_{i-1} \approx 1/3000\) (see § 4.1), it follows that we allow \(k\) to vary only from about \(-150\) to 150. If the minimizing value of \(k\) turns out to be at one of the \(k\) limits, then we consider the test invalid. This occasionally happens and indicates the tested spectra are likely not very similar and likely cannot be made to look similar shifting them relative to each other in logarithmic wavelength: if they can, the resemblance is likely accidental. Also we only consider DIFF2 tests valid if the logarithmic wavelength range overlap when the spectra are not shifted is \(\geq 0.2 \times \log(e)\) which corresponds to a relative wavelength overlap of about 20%. This is the same overlap criterion as used for the DIFF1 test and, as mentioned in § 4.1, is almost always met for spectra we consider.

By its nature, DIFF2 ≤ DIFF1 for any spectrum pair. Thus, DIFF2 also has an upper limit of 2. But it is even less likely for DIFF2 than for DIFF1 that any value will approach 2. To do that would imply that the spectra were anticorrelated no matter how one shifted them over the allowed shifting range. Since even accidental anticorrelations of spectra in a large sample will tend to be evaded by the shifting, we do not expect DIFF2 values much greater than 1.

We have subjected the set of all possible spectrum pairs drawn from the sample of HDCC supernovae described in § 5 to the DIFF2 test. (There are 16 supernovae, 156 spectra, and 21422 valid spectrum pairs: some possible pairs are excluded as discussed below.) The resulting distribution of DIFF2 values has mean, standard deviation, minimum value, and maximum value of, respectively, 0.781, 0.118, 0.229, and 1.131 (where we have reported more digits than are significant to allow for numerical consistency checks): naturally these values are all a bit smaller than the corresponding values for DIFF1 reported in § 4.1. The maximum value as expected is not much greater than 1.

The rationale for DIFF2 is as follows. The structure of P-Cygni line profiles in supernovae is correlated with the velocity structure of the ejecta since that velocity structure determines the Doppler shifts of line features. Most noticeably, weak, unblended lines in the photospheric phase tend to have their absorption features reach a minimum at a wavelength shift corresponding to the photospheric velocity (e.g., Jeffery & Branch 1990, p. 188). All
parts of the absorption feature have blueshifts that are usually correlated with the photospheric velocity in the sense that the greater the photospheric velocity, the greater the blueshift of each part of the absorption feature: this is just because the whole velocity structure of the line-forming region at any phase tends to be correlated with the photospheric velocity at that phase. In comparing photospheric spectra for similar supernovae, there is often some difference in the width scale of line profiles, particularly the absorption features, attributable to some difference in the velocity structure. The difference may be intrinsic to the supernovae: one supernova may just have layers that move faster than the corresponding layers in the other supernova at the same phase. On the other hand, the difference may just be a matter of phase since a photosphere recedes into the ejecta as time passes and the overall optical depth of the supernova falls. Of course, intrinsic and phase differences in velocity structure are usually both present to one degree or another.

The eye can usually discern similar line profile patterns despite differing velocity structure. However, the DIFF1 test may not pick out a striking similarity because it counts vertical differences in the spectra which can be large if the steep edges absorptions and emissions are offset by velocity-structure induced Doppler shifts. One can imagine compensating for the different velocity structures by a wavelength varying shift, but this seems too complex for practical use. Since the blueshifted absorption features are often larger than the emission features (necessarily the case if the lines are pure scattering (e.g., Jeffery & Branch 1990, p. 189)), a general wavelength shift that tends to align the absorption features may bring out similarities both to the eye and in the DIFF2 test (i.e., by minimizing the DIFF2 function). Branch et al. (2006a) showed that general shifts were useful in bringing out similarities. We expect that the DIFF2 test shift will usually try to align blueshifted absorption features in photospheric phase spectra.

In the nebular phase, the emission lines tend to be symmetric about the rest-frame line-center wavelength and DIFF2 will tend to reduce to DIFF1. However, if the emission peaks are offset from the line-center wavelengths because of asymmetry, then the DIFF2 test will compensate for that and again bring out similarities that could be missed by the DIFF1 test.

Because the logarithmic wavelength shift used in DIFF2 will often compensate for differences in velocity structure, it is useful to define a corresponding velocity shift parameter characteristic of the difference in velocity scale between two spectra. Let us say that the wavelength shift $\delta \lambda$ from an initial wavelength $\lambda$ is an actual correction for a Doppler shift. From the first order Doppler formula

$$\frac{\delta \lambda}{\lambda} = -\frac{v}{c},$$

(18)
where we define positive velocities as giving blueshifts. Say the logarithmic wavelength shift is \( \delta \log(\lambda) \) and recall we only allow relative wavelength shifts of up to about 5%. We can then make the approximation

\[
\delta \log(\lambda) = \log(e) \times \delta \ln(\lambda) \approx \log(e) \times \frac{\delta \lambda}{\lambda}
\]

(19)

Thus, it seems appropriate to define the velocity shift parameter corresponding to the logarithmic wavelength shift by the formula

\[
v = -c \frac{\delta \log(\lambda)}{\log(e)} .
\]

(20)

4.3. An Example of Two Spectra Fitted by Minimizing DIFF2

Figure 6 shows an example of two spectra with a relative shift that minimizes DIFF2. In this case, DIFF1 = 0.642 and DIFF2 = 0.612 and the velocity shift parameter corresponding to the shift in logarithmic wavelength is 999 km s\(^{-1}\). The change from DIFF1 to DIFF2 and the velocity shift are not large in this case and, in fact, the eye cannot detect any overall improvement in fit in going from a plot without the shift to the plot with it. The small change between the DIFF1 and DIFF2 and the small shift is to be expected. First, the spectra both come from Type IIb supernovae: the solid-line spectrum is the SN 1993J 1993 April 15 (day −2) spectrum shown in Figure 2 (§3.2) and the dotted-line spectrum is the SN 1987K 1987 August 9 (day 9) spectrum (Filippenko 1988). Second, out of our current and preliminary sample of hydrogen-deficient spectra (see §5), the SN 1987K 1987 August 9 spectrum is the closest match according to DIFF2 to the SN 1993J 1993 April 15 spectrum aside from other spectra from SN 1993J that come from close to 1993 April 15. According to DIFF1, it is only the second closest non-SN-1993J match: the SN 1987K 1987 August 7 is slightly closer with DIFF1 = 0.624.

The phase (day −2) of the SN 1993J spectrum is relative to the well determined UVOIR bolometric maximum light (see §§3 and 3.2). The phase (day 9) of the SN 1987K spectrum is relative to red-optical maximum that is estimated to be 1987 July 31 with an uncertainty of ±4 days (Filippenko 1988). Supernova SN 1993J was a particularly well-observed supernova and hence our precise knowledge of its phase. Supernova SN 1987K was only moderately well-observed and hence our uncertainty about its phase and light curve. The data available for SN 1987K is typical for the supernovae that must dominate any current sample of supernovae for statistical spectral analysis.

The identifications in Figure 6 are the same as in Figure 2 (§3.2), except for the residual telluric absorption lines in the SN 1987K spectrum (Filippenko 1988). The region from the
blue side of the telluric lines to the red end of the SN 1987K spectrum were excluded from the evaluation of DIFF1/2.

4.4. Comparing DIFF1 and DIFF2

Both DIFF1 and DIFF2 measure the similarity of spectra. Because DIFF1 has no compensating logarithmic wavelength shift, a spectrum pair with very small DIFF1 value will likely come from nearly identical supernovae. Because of the compensating logarithmic wavelength shift, DIFF2 should be somewhat better than DIFF1 at finding similarity among supernovae that are physically distinct. In essence, both tests do the same thing, but with a somewhat different weighting.

The most obvious physical distinction that the logarithmic wavelength shift in the DIFF2 test can compensate for is in velocity structure: the distinction is caused either by an intrinsic difference or a phase difference. Another possible distinction is in viewing direction for supernovae that are asymmetric. There is, in fact, considerable evidence from supernova polarimetry and spectropolarimetry for asymmetry in core-collapse supernovae (e.g., Wheeler & Benetti 2000; Wang et al. 2001; Leonard & Filippenko 2005). One important kind of asymmetry can be parameterized by an axial ratio for characteristic orthogonal axes perpendicular to the line of sight. Spectropolarimetric observations suggest that this axis ratio can vary from 1 to 2 or more (e.g., Wang et al. 2001). For example, the Type IIb supernovae SN 1993J and SN 1996cb spectropolarimetric data suggest an axis ratio of $\gtrsim 1.4$ for both these supernovae (Wang et al. 2001). There are data that suggest, but with considerable uncertainty, that asymmetry increases with decreasing hydrogen envelope mass (Wang et al. 2001; Leonard & Filippenko 2005). Another kind of asymmetry in core-collapse supernovae, that of a jet (or bipolar jets) emerging from the main part of the ejecta, is suggested by the spectropolarimetry of Type Ic hypernova SN 2002ap (Kawabata et al. 2002; Leonard et al. 2002). Type Ia supernovae also show polarization in some cases: the polarization may arise from clumps in the ejecta (e.g., Leonard et al. 2005).

Because of the use of local normalization, both DIFF1 and DIFF2 tests are suitable for studying supernova classes with uncertain continua. HDCC supernovae (which often occur in or near star-forming regions) often have uncertain reddening, and so are suitable for studies with DIFF1 and DIFF2.
5. A PRELIMINARY STATISTICAL ANALYSIS OF HYDROGEN-DEFICIENT CORE-COLLAPSE SUPERNOVA SPECTRA

We are beginning a project of comparative analysis of the spectra of HDCC supernovae (i.e., hydrogen-deficient core-collapse supernovae of Types IIb, Ib, Ic, and Ic hypernovae: see § 1) that will include a statistical analysis of the spectra as well a spectral modeling using the parameterized spectrum-synthesis code SYNOW (e.g., Branch et al. 2003, 2005, and references therein). To undertake this analysis project, we are assembling HDCC supernova spectra in html files in supernova directories at SUSPEND under the heading *Supernovae by Epoch*. The spectrum files have header information about the spectra and the supernovae they come from. Plots of the spectra in original form and locally-normalized form are also given in the headers. The original spectra follow the header in two-column format: the columns being wavelength and flux. The locally-normalized spectra are given in the dirs subdirectories of the supernova directories. A list of the spectra are currently in html files (including those spectra from non-HDCC supernovae) can be found under the heading *Lists* at SUSPEND.

At present, we are at a very early stage in the project and have not assembled many of the spectra available for HDCC supernovae nor determined the best way to proceed with our statistical analysis. However, to illustrate the use of DIFF1 and DIFF2, we present some preliminary results.

5.1. Table 1 and Tables at SUSPEND

Table 1 gives some statistics on the supernovae and spectra we have assembled so far in SUSPEND. The supernovae include the prototype Type Ib’s SN 1983N and SN 1984L (Harkness et al. 1987) and the well-observed Type Ic’s SN 1987M (Filippenko et al. 1990) and SN1994I (e.g., Clocchiatti et al. 1996). The peculiar Type Ibc SN 2005bf (e.g., Folatelli et al. 2006) is included as well: this supernova had two post-early-phase maxima, had a risetime of about 40 days to a second and main optical maximum (typical Type Ib/c risetimes are of order 20 days), and was unusually luminous. It also made a transition from Type Ic to Type Ib behavior in the period from about 10 days to about 60 days after explosion: because of this transition, it is classified as a Type Ibc. (Type Ibc supernovae are those that are not clearly distinguishable into Type Ib’s or Type Ic’s.) Only two Type IIb supernovae have been included so far: the already discussed prototype SN 1987K (§ 4.3; Filippenko 1988) and the very well-observed SN 1993J (§ 3 and § 4.3; e.g., Jeffery et al. 1994; Richmond et al. 1994; Baron et al. 1995; Clocchiatti et al. 1995; Houck & Fransson 1996; Fransson et al. 2005). Three Type Ic hypernovae included: SN 1997ef (e.g., Mazzali et al. 2000), SN 1998bw (e.g.,
Galama et al. 1998; Iwamoto et al. 1998; Patat et al. 2001), and SN 2002ap (e.g., Kawabata et al. 2002; Leonard et al. 2002; Foley et al. 2003). Also included are less famous HDCC supernovae observed by Matheson et al. (2001) and others.

In the calculations described below, we include only one spectrum for each phase period (which we usually set to being 1 day) from each supernova. The other spectra for that phase period are redundant since they should be nearly identical to the included spectrum aside from observational error. The spectrum we include is the one we deem to be best: this is usually because of broader wavelength coverage.

In an automated fashion, we have calculated DIFF1 and DIFF2 for all supernova spectrum pairs in the sample: these spectrum pair DIFF1/2 values can be found tabulated at SUSPEND under the heading Lists. For each spectrum, there is a table listing all other spectra in order of increasing DIFF1/DIFF2 values relative to it. These tables are updated as more supernovae and spectra are added to the sample. We have excluded from the analysis below and in § 5.2 those DIFF1 and DIFF2 values deemed invalid by the rules discussed in §§ 4.1 and 4.2.

As well as spectra, we wanted some way to order supernovae in order of likeness. As a preliminary method, for any two supernovae we calculate all (valid) DIFF1/2 values between spectrum pairs with one of the pair coming from one supernova and the other from the other supernova: i.e., we calculate the non-self pair DIFF1/2 values. The smallest of these DIFF1/2 values we call the supernova pair DIFF1/2 value. The rationale for this procedure for supernova pair DIFF1/2 values is that the spectral phase coverage for all supernovae is incomplete (although for some the coverage is very good) and rarely are both supernovae covered at an exact common phase: often coverage at even a nearly common phase is lacking. Also phase relative to maximum light is itself uncertain. For some supernovae, one has accurate dates for the UVOIR bolometric maximum light or some broadband maximum light. For other supernovae, one only has some estimate of what can be loosely called optical maximum light which should approximate UVOIR bolometric maximum light to some degree. In some cases, the estimate is just the day of discovery: supernovae tend to be discovered near maximum light since that is when they are most readily discoverable because, of course, they are brightest at maximum light. For the assignment of maximum light, we use UVOIR bolometric maximum light when available (which is rarely) and what we believe to be the best surrogate for the UVOIR bolometric maximum light in other cases. By choosing the smallest DIFF1/2 value out of all the non-self pair values for the supernova pair DIFF1/2 value, we hope to have partially evaded the problem of incomplete spectral phase coverage. We assume that the likeness of the supernovae is best represented by their closest observed spectral approach to each other. For supernovae that are actually much alike, this closest
approach is probably when the spectrum pair are from nearly the same phase, and thus the spectrum pair gives the best determination of actual likeness of the supernovae.

The tables giving the spectrum pair DIFF1/2 values at SUSPEND (which altogether are quite lengthy) are preceded by tables of supernova pair DIFF1/2 values.

5.2. Table 2 and the Standard HDCC Supernova Types

Since this paper is just a beginning in our project of statistical analysis of HDCC supernova spectra, we will not discuss the significance of either the tabulated supernova or spectrum pair DIFF1/2 values at SUSPEND (see § 5.1). But as a preliminary investigation, we consider if the classification of HDCC supernovae into Type IIb’s, Type Ib’s, Type Ic’s, and Type Ic hypernovae is confirmed by the supernova pair DIFF1/2 values. (Supernova classified as Type Ibc (of which there is only one, SN 2005bf, in the current sample) are grouped with the Type Ib’s.) We have therefore calculated the mean supernova pair DIFF1/2 values between supernovae of different types (the cross-type means) and the mean supernova pair DIFF1/2 values between supernovae of the same type (the self-type means). We have also calculated the estimated standard deviations of the distributions of DIFF1/2 values using the ordinary standard deviation formula with the correction for using the mean of the sample (e.g., Bevington 1969, p. 19). The standard deviation for the Type IIb pairs is assigned a zero value since there are only two Type IIb’s in the current sample and, thus, only one pair value: the standard deviation of the sample is zero; that of the distribution is unknown. Since HDCC supernovae tend to look increasingly alike as time from explosion (i.e., phase) increases, we have repeated the DIFF1/2 calculations including only supernova pair DIFF1/2 values for spectra that are both from phases less than 10 days past maximum light.

The results of all the calculations are given in Table 2 in the form of a type correlation matrix consisting of self-type means (the diagonal elements) and cross-type means (the off-diagonal elements): phase-unrestricted DIFF1 values are in Table 2(a), phase-unrestricted DIFF2 values are Table 2(b), phase-restricted DIFF1 values are in Table 2(c), and phase-restricted DIFF2 values are in Table 2(d). The matrix is, of course, symmetric.

Several remarks can be made about the results in Table 2. First, none of the mean values are very small: they are all larger than 0.455 which is the Table 2(b) Type Ic DIFF2 self-type mean. An examination of the table of supernova pair DIFF1/2 values at SUSPEND shows that none are less 0.4: the smallest DIFF1 is 0.437 for the pair of Type Ib’s SN 1999dn and SN 19994l; the smallest DIFF2 is 0.420 also for the same pair of supernovae and using
the spectrum pair. Some of the spectrum pair DIFF1/2 values are significantly smaller (going down to 0.244 for DIFF1 and 0.229 for DIFF2), but those are for spectra from same supernova usually taken close together in phase. The upper limit on the phase-unrestricted supernova pair DIFF1 values is 0.886 and on the phase-unrestricted supernova pair DIFF2 values is 0.809.

Second, a DIFF2 value is always less than or equal to the corresponding DIFF1 value as it must be according to our formulae (§ 4.1, eq. (12) and § 4.2, eq. (17)), but always by less than the standard deviation of the DIFF1 value. We have to conclude that DIFF2 is not finding great similarities that are being totally missed by DIFF1.

Third, the phase-restricted DIFF1/2 values (in Table 2(c) and Table 2(d)) are usually bigger than the corresponding phase-unrestricted (in Table 2(a) and (b)). This is understandable since HDCC supernovae are qualitatively understood to be more diverse at earlier times. However, one may also be significantly reducing the relative phase coverage for some supernovae by the phase restriction if they were poorly sampled at early times. The two effects would need to be sorted out in a fuller analysis. The exceptions to the usual case are the Type Ic self-type means that actually decrease going from the phase-unrestricted to the phase-restricted cases. The reason for this decrease is simple: one of the 4 Type Ic’s in the sample, SN 1990aa, has no spectra available for 10 days or less past maximum light. In fact, we only have one spectrum for SN 1990aa (from day 17 past visible maximum light) in the current sample: other spectra exist and will be included eventually, but they are, in fact, all from later phases (Matheson et al. 2001). The other Type Ic supernovae in the sample do have spectra from a comparable phase to that of the SN 1990aa, and so SN 1990aa’s divergence from the other Type Ic’s (as evidenced by the increase in the self-mean in going from the phase-unrestricted case to the phase-restricted case) is not likely due to lack of phase coverage. It may be that SN 1990aa is a bit of an outlier in the sample of Type Ic’s and not including it the phase-restricted sample narrows the dispersion of Type Ic’s among themselves.

Fourth, the self-type means in any row are the smallest means in that row in all cases. This confirms the conventional types. In the rows for Type IIb’s and Type Ic hypernovae, the self-type means are smaller than the row cross-type means by $\gtrsim 1.5\sigma_{\text{larger}}$ in all cases and by $\gtrsim 2\sigma_{\text{larger}}$ in all but two cases. (The $\sigma_{\text{larger}}$ is the larger of the standard deviations of any two means under comparison.) Thus, Type IIb’s and Type Ic hypernovae are found to be quite distinct from the other types.

On the other hand, for the phase-unrestricted cases, in the rows for Type Ib/Ibc’s and Type Ic’s the self-type means are smaller than the row cross-type means by $< \sigma_{\text{larger}}$. For the phase-restricted cases, the Type Ib/Ibc self-type means are smaller than the row cross-type
means by $> \sigma_{\text{larger}}$, except for the cross-type mean for between the Type Ib/bc’s and the Type Ic’s for which the difference is $< \sigma_{\text{larger}}$. For the phase-restricted cases, the Type Ic self-type means are smaller than the row cross-type means by $> \sigma_{\text{larger}}$. Taken at face value, these observations suggest that Type Ib/Ibc’s and Type Ic’s are more heterogeneous types than Type IIb’s and Type Ic hypernovae and overlap more in behavior with other types, but restricting DIFF1/2 tests to early phases allows Type Ib/Ibc’s and Type Ic’s to be more clearly distinguished. These conclusions must be tentative, however, given the incompleteness of the current sample.

The Type IIb’s and Type Ic hypernovae included in the sample are all well-observed supernovae, and so they suffer relatively little from lack of phase coverage. Some of the Type Ib/Ibc’s and Type Ic’s in the sample have very poor phase coverage. (See the list of supernova spectra in html format under the heading *Lists* at SUSPEND for the spectra currently available for individual supernovae.) The supernova pair DIFF1/2 values for supernovae where both supernovae have poor phase coverage could be large even if the supernovae were very similar simply because the procedure for determining the supernova pair DIFF1/2 values may not be able to find spectra from nearly the same phase. In future work, we will try to remedy this deficiency in the procedure perhaps by giving lower weight to supernova pair DIFF1/2 values from supernova pairs where both members of the pair have poor phase coverage or by using some quite different definition of supernova pair DIFF1/2.

6. CONCLUSIONS AND DISCUSSION

We have developed two tests DIFF1 and DIFF2 for measuring goodness-of-fit between two supernova spectra (see § 4.1, eq. (12) and § 4.2, eq. (17) for the formulae). The tests rely on local normalization which eliminates the problem of uncertainty in the spectrum continuum (§ 2). It also eliminates real information stored in continuum shape. However, our basic premise is that line pattern is a much better signature of intrinsic supernova behavior than continuum shape, and so eliminating continuum shape information is not too important.

We have presented some examples of locally-normalized spectra for SN 1993J and given some analysis of the spectra (§ 3). Two of the SN 1993J spectra are hitherto unpublished *HST* UV spectra. We have shown that local normalization is reliable in regard to continuum distortions of the original spectrum: i.e., the locally-normalized spectra one obtains from various distorted versions of the original spectrum are nearly identical as judged by eye and the DIFF1 test. One must, however, apply the same local normalization procedure in all cases for this reliability.
As an example of the use of DIFF1/2, we have used them in a preliminary statistical analysis of the spectra of HDCC (hydrogen-deficient core-collapse) supernovae. The analysis confirms that conventional HDCC supernova types (IIb, Ib/Icon, Ic, and Ic hypernova) do form distinct groups when compared using DIFF1/2. This analysis is preliminary since many available HDCC supernova spectra are not included in the analysis sample. Also many improvements in our statistical procedure are possible.

Preliminary statistics tables for our sample of HDCC supernovae are available at the SUSPEND database (see the footnote to the abstract for the URL) under the heading Lists. The spectra we have used are also online at SUSPEND in two-column format under the heading Supernovae by Epoch in supernova directories in html files along with figures of the original and locally-normalized spectra. The locally-normalized versions of the spectra themselves can be found in the dif subdirectories of the supernova directories.

Although hundreds of supernovae are now being discovered per year (e.g., 367 in 2005 (Central Bureau for Astronomical Telegrams 2006) which is more than 1 per day), most of these are remote and are relatively poorly observed. Their main use is as cosmological probes and for the determination of supernova rates. New well-observed supernovae accumulate slowly with only a few per year. These new well-observed ones and the past well-observed ones are only relatively well-observed in most cases. Phases are missing, calibrations imperfect, and frequently the reddening correction is very uncertain. Thus, for the foreseeable future, statistical analyses of supernova spectra will have to rely on heterogeneous data sets for relatively few well-observed supernovae. We believe statistical analyses with DIFF1/2 (which eliminate the need for accurate continuum shape and that can compensate for varying phase, velocity structure, and asymmetry) will be a useful tool in statistical analyses of available spectra.

Another, and not-distinct, use for DIFF1/2 is to find for new and, perhaps, not-well-observed supernovae, a well-observed supernova that is a near-twin. Then insofar as the well-observed supernova is understood, the new supernovae will be understood.

DIFF1/2 should also be useful in synthetic spectrum analysis of supernova spectra because, again, it eliminates the need for dealing with the uncertain continuum of observed supernovae. Certainly fitting the true continuum is one of the goals and guides in achieving a good synthetic spectrum fit to the observations. But when the fitting exercise gives a good fit to an incorrect continuum, then the fitting exercise becomes misleading. A good fit to the lines with highly realistic radiative transfer and a realistic hydrodynamic model should yield a good continuum and that should allow one to correct the observed continuum for calibration errors and reddening.
The future of synthetic spectrum modeling of supernovae may well be the calculation of time-dependent radiative transfer for a large collection of realistic hydrodynamic models that span many possible explosion outcomes. As time passes more models will be included in the collection and the realism of the radiative transfer will be improved. Eventually, all old and new supernovae will find a near-twin in the collection using some test like DIFF1/2 and will then be well understood. This ideal situation is still far off, but there is work heading toward it (e.g., Kasen 2006, for Type Ia supernovae).

In the near future, we plan to use DIFF1/2 in a statistical analysis of HDCC supernova spectra. This analysis will also include spectral modeling using the parameterized spectrum-synthesis code SYNOW (e.g., Branch et al. 2003, 2005, and references therein). Use of DIFF1/2 for other supernova types is also envisaged for future work.

We would like to thank all the observers who obtained the spectra used in this paper and Peter Challis who prepared the two nebular-phase HST spectra for us some time ago. Support for this work has been provided by NASA grant NAG5-3505 and the Homer L. Dodge Department of Physics & Astronomy of the University of Oklahoma.

REFERENCES


Branch, D., et al., 2006b, PASP, 118, 560

Buta, R. J. 1982, PASP, 94, 578


Challis, P. 1994, private communication


Filippenko, A. V. 1994, private communication


Fransson, C. 1984a, 132, 115
Fransson, C. 1984b, 133, 264


INES (IUE Newly Extracted Spectra) 2006, The INES Archive Data Server: IUE Final Archive (Villafranca del Castillo, Spain: Laboratory for Space Astrophysics and Theoretical Physics), http://iuearc.vilspa.esa.es/ines/index2.html


Kasen, D. 2006, in preparation


NIST (National Institute of Standards and Technology) 2006, NIST Atomic Spectra Database Version 3.0.3 (Gaithersberg, Maryland: NIST), http://physics.nist.gov/PhysRefData/ASD/lines_form.html


Ripero, J. 1993, IAU Circ. 5731


Wamsteker, W., Rodriguez, P. M., Gonzalez, R., Sonneborn, G., & Kirshner, R. P. 1993, IAU Circ. 5738


Fig. 1.— The SN 1993J (Type IIb) spectrum in the $f_\lambda$ and locally-normalized representations from 1993 March 31 which is about 16 days before UVOIR bolometric maximum light and about 4 days after explosion. The locally-normalized spectra in this and in other figures are obviously the ones with continuum level of about 1.
Fig. 2.— The SN 1993J (Type IIb) spectrum in the $f_{\lambda}$ and locally-normalized representations from 1993 April 15 which is about 2 days before UVOIR bolometric maximum light and 18 days after explosion. From 3240Å blueward, the spectrum is an HST spectrum.
Fig. 3.— The SN 1993J (Type IIb) spectrum in the $f_\lambda$ and locally-normalized representations from 1993 September 15.5 which is about 151 days after UVOIR bolometric maximum light and 171 days after explosion. From 4253.6 Å blueward, the spectrum is an $HST$ spectrum (from 1993 September 17).
Fig. 4.— The SN 1993J (Type IIb) spectrum in the $f_\lambda$ and locally-normalized representations from 1993 November 15.5 which is about 212 days after UVOIR bolometric maximum light and 232 days after explosion. From 4200 Å blueward, the spectrum is an HST spectrum (from 1993 November 14). The original HST spectrum blueward of $\sim 2425$ Å seems to decline in a manner too steep to be physically real and we do not trust it nor the locally-normalized spectrum there.
Fig. 5.— The evolution of the blueshifted Mg II λ2797.9 emission line in the SN 1993J spectra. The phases given in parentheses in the figure are relative to the UVOIR bolometric maximum light on 1993 April 17.
Fig. 6.— The locally-normalized spectra of Type IIb supernovae SN 1993J from 1993 April 15 (about 2 days before UVOIR bolometric maximum light) and SN 1987K from 1987 August 9 (about 9 days after optical maximum light). The SN 1987K spectrum has been blueshifted by velocity 999 km s$^{-1}$ to minimize the DIFF2 value for the spectrum pair. For the spectrum pair, DIFF1 = 0.642 and DIFF2 = 0.612. The spectra were both initially corrected for host galaxy heliocentric velocity: from Leda, the mean host galaxy heliocentric velocities are $-39 \pm 2$ km s$^{-1}$ (SN 1993J) and $799 \pm 2$ km s$^{-1}$ (SN 1987K) (Paturel et al. 2003).
Table 1. Current Sample in SUSPEND of HDCC Supernovae and Spectra

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of Supernovae</th>
<th>Number of Spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>All HDCC Types</td>
<td>16</td>
<td>156</td>
</tr>
<tr>
<td>Type IIb</td>
<td>2</td>
<td>39</td>
</tr>
<tr>
<td>Type Ib/Ibc&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>Type Ic</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>Type Ic hypernova</td>
<td>3</td>
<td>50</td>
</tr>
</tbody>
</table>

<sup>a</sup>Type Ib<sub>c</sub> supernovae are those supernovae which are not clearly distinguishable into Type Ib’s or Type Ic’s. We group Type Ib’s and Type Ibc together. There is only one Type Ibc in the sample: SN 2005bf.
Table 2. HDCC Supernova Type DIFF1/2 Correlation Matrix

<table>
<thead>
<tr>
<th>Type\Type</th>
<th>Type IIb</th>
<th>Type Ib/Ibc$^a$</th>
<th>Type Ic</th>
<th>Type Ic hyp$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean DIFF1/2</td>
<td>St.Dev.</td>
<td>Mean DIFF1/2</td>
<td>St.Dev.</td>
</tr>
<tr>
<td>(a) DIFF1: no phase restriction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type IIb</td>
<td>0.507</td>
<td>0.000</td>
<td>0.676</td>
<td>0.083</td>
</tr>
<tr>
<td>Type Ib/Ibc</td>
<td>0.676</td>
<td>0.083</td>
<td>0.624</td>
<td>0.097</td>
</tr>
<tr>
<td>Type Ic</td>
<td>0.687</td>
<td>0.077</td>
<td>0.662</td>
<td>0.109</td>
</tr>
<tr>
<td>Type Ic hyp</td>
<td>0.632</td>
<td>0.052</td>
<td>0.709</td>
<td>0.082</td>
</tr>
<tr>
<td>(b) DIFF2: no phase restriction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type IIb</td>
<td>0.507</td>
<td>0.000</td>
<td>0.643</td>
<td>0.076</td>
</tr>
<tr>
<td>Type Ib/Ibc</td>
<td>0.643</td>
<td>0.076</td>
<td>0.573</td>
<td>0.096</td>
</tr>
<tr>
<td>Type Ic</td>
<td>0.632</td>
<td>0.038</td>
<td>0.612</td>
<td>0.105</td>
</tr>
<tr>
<td>Type Ic hyp</td>
<td>0.601</td>
<td>0.042</td>
<td>0.668</td>
<td>0.060</td>
</tr>
<tr>
<td>(c) DIFF1: restricted to pre-10-days past maximum light</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type IIb</td>
<td>0.507</td>
<td>0.000</td>
<td>0.732</td>
<td>0.068</td>
</tr>
<tr>
<td>Type Ib/Ibc</td>
<td>0.732</td>
<td>0.068</td>
<td>0.632</td>
<td>0.076</td>
</tr>
<tr>
<td>Type Ic</td>
<td>0.731</td>
<td>0.089</td>
<td>0.674</td>
<td>0.091</td>
</tr>
<tr>
<td>Type Ic hyp</td>
<td>0.735</td>
<td>0.086</td>
<td>0.855</td>
<td>0.130</td>
</tr>
<tr>
<td>(d) DIFF2: restricted to pre-10-days past maximum light</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type IIb</td>
<td>0.507</td>
<td>0.000</td>
<td>0.695</td>
<td>0.058</td>
</tr>
<tr>
<td>Type Ib/Ibc</td>
<td>0.695</td>
<td>0.058</td>
<td>0.580</td>
<td>0.062</td>
</tr>
<tr>
<td>Type Ic</td>
<td>0.658</td>
<td>0.049</td>
<td>0.633</td>
<td>0.087</td>
</tr>
<tr>
<td>Type Ic hyp</td>
<td>0.681</td>
<td>0.069</td>
<td>0.769</td>
<td>0.077</td>
</tr>
</tbody>
</table>

$^a$Type Ib supernovae are those supernovae which are not clearly distinguishable into Type Ib’s or Type Ic’s. We group the Type Ib’s and Type Ib’s together. There is only one Type Ib in the sample: SN 2005bf.

$^b$Type Ic hyp is short for Type Ic hypernovae.

Note. — This table gives the mean DIFF1/2 values between HDCC supernova types in a correlation matrix format. The DIFF1/2 value between two supernovae in this preliminary analysis is just defined to be the smallest DIFF1/2 value found out of all the pairs of spectra with one spectrum drawn from the one supernova and the other from the other supernova. With no phase restrictions the spectra are drawn from all phases. With phase restriction, the spectra are only drawn from the given phases. The matrix is symmetric, of course. The diagonal elements are the mean DIFF1/2 values for supernova pairs of the same type: we call these means the self-type means. The off-diagonal elements are the mean DIFF1/2 values for supernova pairs of different types: we call these means the cross-type means. More digits are shown than are significant to allow for numerical consistency checks.