TIME DOMAIN CHARACTERISATION OF MODERN TIME AND FREQUENCY STANDARDS

Richard Percival
Quartzlock (UK) Ltd
Gothic
Plymouth Road
Totnes, Devon, TQ9 5LH
0044 (0)1803 862062
richard@quartzlock.freeserve.co.uk

Abstract - The article is an attempt at explaining the basic theory and practice behind frequency calibration in the time domain. There are many different ways of characterising the frequency accuracy and stability of a standard in the time domain. However, successful characterisation involves, amongst other things, consideration of the various noise processes generated by frequency standards, understanding the multitude of statistical methods available, following a quality calibration procedure (such as NAMAS) and the efficient logging of the results. This paper seeks to explain and clarify the terminology used to specify and characterise modern Time and Frequency Standards. It shows the difference between accuracy and stability, expounds the need for appropriate reference standards, charts the development of the Allen Variance and its cousins, gives detailed information into their respective advantages and disadvantages and shows where each may be most appropriately used. To aid in understanding how time domain measurements are taken, the reader is led through the procedure used in-house by the company to characterise and calibrate their range of time and frequency standards. This is both for internal (research and development) and external (calibration certification for customers) use.

This document is an attempt to show procedures used to calibrate and characterise time and frequency standards for both commercial and research and development purposes.

Quartzlock is a manufacturer accredited to ISO 9001 standard. This means that we are able to fully document every stage of our working practice. This extends to the testing done on our range of frequency generation, distribution and measurement instruments.

The most important measure of performance for frequency standards is the frequency stability. The standard to be calibrated is labelled the device under test (DUT). In order that this DUT be calibrated and characterised, it must be compared to a frequency standard of higher performance, generally called the reference standard. As detailed in an excellent article by Michael A. Lombardi of NIST **[1]**, in order that the testing be valid the reference standard must, at all times during this test, outperform the DUT by a specified ratio, known as the test uncertainty ratio (TUR). The ratio itself is not explicitly mentioned, although a TUR of 10:1 is generally an acceptable ratio. This ensures that the time taken for the DUT to be unambiguously tested is reasonable. It is, of course, possible for a TUR of 5:1 to be used. However this can lead to, an albeit small, error contribution of the reference standard into the phase residuals collected, which will lead to inaccuracy in the calibration results.

Quartzlock are manufacturers of a broad range of frequency standards including lf tracking receivers, OCXO's, Hydrogen Masers, Rubidium atomic frequency standards and GPS/GLONASS receivers. In order that the required TUR ratio of 10:1 is maintained at all times during the testing, the company has employed a Passive Hydrogen Maser, as it's in-house reference standard. This ensures that all products can be tested – apart from its masers- to the accuracy required. This includes the new 1999 temperature controlled GPS disciplined Rubidium, with its excellent performance in the short, medium and long term.

There are basically two parameters of interest to the customer and to the company

i) How accurately does the supplied frequency standard agree with the frequency specified (Lombardi calls this the *'nameplate frequency'*). This indicates the ability of the standard to deliver what it promises and is commonly called either the *'accuracy'* or the *'fractional frequency offset'*. There are various methods through which the *'accuracy'* of a frequency standard may be specified. It will be important to understand what is being said in these specifications if misunderstandings are to be avoided. Later on, I will detail how the '*accuracy'* of a frequency standard may be calculated and explain the different methods through which those results may be expressed. It is really an indication of how well adjusted the oscillator is. It indicates coarsely whether the oscillator is subject to drift or aging to any significant degree. It does not, however, impart any information as to the inherent quality of an oscillator. It is a valuable method through which crystal oscillators are specified and instrumentation for use in basic calibration is compared.

ii) How *'Stable'* is the output signal from the frequency standard over a period of time. Whilst this does not indicate the 'rightness or wrongness' of a signal it does indicate whether it changes. It is generally accepted that the stability of a frequency standard will (usually) not change when the frequency offset is removed through calibration. Lombardi defines '*Stability'* as *'the statistical estimate of the frequency fluctuations of a signal over a given period of time'*. It is the *'Stability'*, which indicates how good the oscillator is. Whilst it is possible to calculate the frequency stability in either the time or the frequency domain, it is generally performed in the time domain. If stability data is required for averaging times of <1s then the data must be gathered in the frequency domain and then a conversion made between the two domains **[2].**

It is important to realise that there are many different ways of characterising the frequency stability of a standard in the time domain. As detailed by Allan et al **[2],** the major problem with frequency standards is that they *'do not generate a constant frequency output contaminated solely by white noise*'. If they did we could average the output to get rid of the noise. That is, white noise is one kind of noise that can be averaged away. This is because for every '*phase advance it produces in the output of the frequency standard, it eventually produces a compensating phase retardation, so that the two cancel in the averaging process'*. This prevents the use of the standard deviation to statistically measure the set of data points. This is because, however many data points we collect, we will never approach the true mean. Indeed the more data we collect the further away we are likely to get from the mean. It has been identified that there are many other types of noise, which afflict frequency standards **[2][5]**. Among those discussed in the 2 referenced papers are white noise PM (α=+2)-phase modulation, flicker PM (α=+1), white noise FM (α=0)-frequency modulation, flicker FM (α=-1) and random walk FM (α=-2). To understand the origins of these noise processes the following may help; the phase modulation noise processes are likely to be related to internal characteristics of the devices themselves, whereas the frequency modulation noise has it's origins in environmental factors, such as temperature variations, mechanical shock and path delay variations.

As a result, new variances had to be developed, which are able to take into account these different types of noise **[6].** All of the newly developed variances were based around the need to filter out the different trends associated with the outputs of atomic frequency standards. To accomplish this, a method of taking the '*first differences*' was used, whereby for all data points, the kth data point is taken away from the k th+1 data point, thereby removing any long term trends in the data (to do with ageing and drift). As Allan et al **[2]** describe, this method has the same effect as a '*digital filter in the frequency domain whose characteristics can be defined precisely*'. By repeating this process twice, we are able to remove nearly all of the noise processes that are encountered in modern frequency standards and clocks. So successful was this differencing approach, that it is still the mainstay of all time domain approaches to the characterisation of time and frequency systems. However, to accurately characterise different types of systems, where the quantity of interest (i.e. time or frequency etc) differs, several different variances have been developed. Essentially all are mere modifications of the original Allen Variance **[6]**

Before detailing the different variances now available to the modern metrologist, it is necessary to understand how the data points are taken from the DUT. To do this, I will use the measurement system in use at Quartzlock both as a general example and also to show how we do our in-house calibrations.

Quartzlock utilise their own A7 frequency/phase difference comparator, which consists of a 2U rack unit with the inputs and controls, a PC card counter which must be installed in a suitable personal computer and the data analysis package Stable32™.

The comparator will operate at either 5 MHz or 10 MHz with automatic switching. The inputs are 50 impedance, and a level of between 6 dBm and 13 dBm is required at both inputs. The absolute accuracy of both reference and measurement inputs should be less than +/-1E-7. The maximum frequency difference should be less than +/-1E-8 in frequency mode and 1E-9 in phase difference mode.

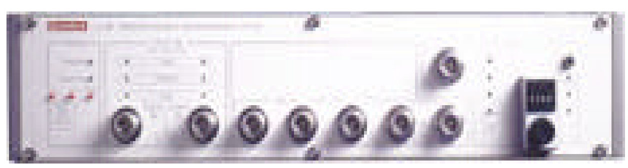

Fig 1: Photo of A7

To collect data for later analysis, the A7 is operated in phase/time difference mode. Here a time interval measurement is taken between a 1 Hz pulse derived by division from the measurement input channel A and the reference 1 Hz pulse derived by division from the reference input channel B. The user who requires information on the precise operation of the Kvarz frequency difference multiplier is advised to refer to **[7].** The counter card use the latest method of time interpolation **[9]** counting to increase the single shot resolution of the A7 to 300 femtoseconds.

An alternative method of data collection is to compare the 1pps output of the DUT with that from a known reference standard. A good reference for 1pps output would again be the hydrogen maser. Because the user is interested in the changes in the time difference (i.e. frequency difference) between the DUT and the reference, it is not important to reference the pulses to GPS/UTC. Whilst the pulses from the maser may well be offset from zero by a small amount, their periodicity is unaffected. This is a classic method to evaluate the performance of GPS-DO's. However, the user should be aware that collection of data by means of different methods leads to different results. The source from which the data was collected should be made explicit when stating the results. The results can often differ by an order of magnitude. In a recent study to determine the suitability of GPS disciplined oscillators as time and frequency standards traceable to the UK national time scale UTC (NPL) **[11]**, one Quartzlock product, the A8-Rb, had an Allan Variance @1000s of 5E-13 from phase data and 1E-12 from the 1pps data. Quartzlock calibrate and characterise solely from phase data. We believe this offers the best method for characterising the outputs from frequency standards.

Before the characterisation of any frequency standards may take place, it is important to verify the performance of the measurement system being used and to characterise the noise floor of the system. It is vital that the noise floor of the system is well below that of any of the frequency standard to be used. If this is not the case there will exist a noise contribution from the measurement system in the DUT characterisation, thereby contaminating the results and casting doubt on the calibration.

Quartzlock operate the following procedure to establish the A7 noise floor and long term drift **[7].** The unit is operated with as near as possible identical signals applied to the reference and measurement inputs to the A7. This is achieved by driving the inputs from two outputs of an inductive power splitter. This ensures that the outputs are identical, and that the noise of the source or distribution amplifier is not being measured. The splitter used by Quartzlock is the Minicircuits ZFSC-2-4 with type N connectors. The source used is generally a Rubidium or a Hydrogen Maser, to ensure that the maximum absolute offset specification of 1E8 is not exceeded (although an OCXO would suffice). The phase noise of the source will actually have some effect and will be transferred directly (without multiplication) to the output of the A7. A 1Hz shift in the absolute frequency of the source will result in a 1 Hz shift in the absolute frequency of the 5 MHz output of the A7. The counter reference will also shift an equal fractional frequency and so no long-term error will occur. However, if the source has very poor short-term stability this may degrade the apparent noise floor of the A7. The unit is then characterised by Allen Variance plot (more on this later) with averaging times from 0.01s to as long as possible. The slope of the curve will give a limit as to long term frequency measurements. For example with two Hydrogen Masers with a 1E-15-frequency difference, the drift rate due to the frequency difference will be 1 femtosecond per second or 3.6 ps per hour. Thus knowledge and control of the measurement system drift rate is of vital importance if very small frequency differences are to be measured

It is important to remember that Allen Variance cannot be calculated for τ less than the time interval between readings (this determines τ_{min}). Also as Allen variance is a statistical measure, τ should not be greater than the total-run-length/10 (this determines τ_{max}). Therefore if 10000 reading are taken each at an interval of 1s, then an Allen Variance plot from 1s to 1000s may be obtained. Ideally the plot should be long enough to see changes in the noise processes occurring and observation of the flicker floor is recommended

Once the measurement system is suitably warm (12 hours is the minimum) and characterised, you are ready to begin the comparison between the DUT and the reference. The DUT has been allowed to warm up for at 24 hours and is connected to the measurement input via a type N connector and the signal from the maser is connected via another type N connector to the reference.

Checks are made that the input levels are suitable (between 6 dBm and 13 dBm). The A7 is then put into phase time difference mode **[7]** and visual checks are made on screen to ensure everything is connected correctly and that the readings seem sensible. It is vital that, whilst being impartial to the results, the metrologist knows 'roughly' what he should be expecting. Otherwise faulty connections etc. may be allowed to go undetected and a run made. This will waste both time and money.

The user then decides over what range he wishes to characterise the unit. The minimum τ is 1s in time difference mode. If the time interval between measurements (referred to as arming) is left at 1 second, then the maximum τ for the Allan Variance Plot will be 3200. This is because of the present limits in memory in the computer used. However, extending the arming value to say 10s, increases τ_{max} to 32000, but also increases τ_{min} by the same amount. Therefore to fully characterise a frequency standard from 1 to 1,000,000 (i.e.10^6 s) two separate runs are necessary. Quartzlock characterises all standards up to $\tau = 10^4$ seconds. It should be remembered that to characterise a frequency standard any longer than this, takes a long time, often unavailable to a busy calibration laboratory. Whilst the A7 has amazing resolution and stability, it is only capable of calibration of 1 device at a time (without an auxiliary switch)

When the run has finished, the raw un-processed data is saved as ASCII data to a floppy disk and transferred to a separate computer for post processing and analysis. The data collection is used solely for this purpose as outlined in NAMAS requirements (Quartzlock is currently seeking accreditation for its calibration and characterisation procedures). Whilst there are limited methods for collection of raw data, there exist a wide range of statistical techniques designed to understand this data, and hence the device tested. Some of these have already been mentioned briefly, and I now intend to explore these in more depth.

There exist three basic variances that are all based upon the differencing method pioneered by D.W. Allan in 1966 [6]. They clearly reveals the types and levels of noise in a particular application '**[6]**

AVAR, or Allan Variance, was developed first and was designed to reveal non-white noise processes present in frequency standards. It is ideal for looking at systems whereby the frequency output is the primary item of interest (i.e. a frequency standard, not a clock). It is of great use in the characterisation of medium to long-term stability of clocks and oscillators and is quite fast to compute. AVAR is the CCIR accepted definition of stability and will be the type used on nearly all specifications and data sheets. It is therefore useful to gain an understanding of how to interpret an AVAR plot. It is defined as below

$$s_y^2(t) = (1/2t^2)<(D^2x)^2>$$

Where $\mathbf{D^2}$ is the second finite difference operator

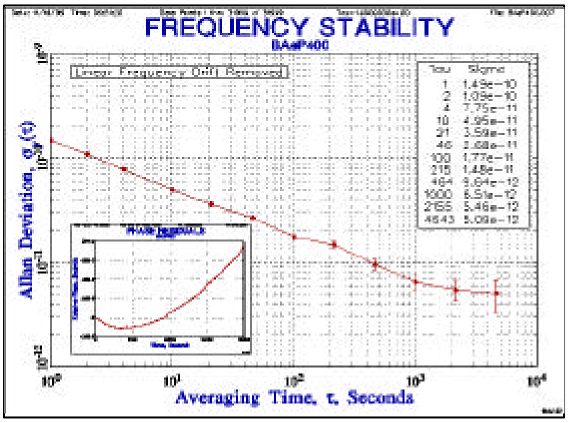

Fig 2: An Allan deviation plot

are AVAR, MVAR and TVAR. It is important that the user understands the applicability of these statistical measures, if one if to make full use of them in differing circumstances. That is, we must choose the appropriate measure 'which most **MVAR**, or modified Allan variance, was developed to overcome a problem encountered with AVAR. AVAR is unable to distinguish between white and flicker noise PM. This is the type of noise related to internal characteristics of the device (such as deformation of a quartz crystal or wall shift problem in active hydrogen maser) and is generally present for short averaging times. Active Hydrogen Masers display white noise PM but not flicker noise PM, whereas Quartz displays both. Unless MVAR has been developed, this would have remained undetected. MVAR conquers this problem by taking the averages of the phase, which changes the bandwidth in the software in just the right way. Again MVAR is very useful for characterisation of frequency standards, especially in the short term <100s, where phase modulation noise is present. This makes it 'suited for electronically generated noise processes' **[2]**

$$s_y^2(t) = (1/2t^2)<(D^2\underline{x})^2>$$

Where x denotes the phase averages used in the second differencing process

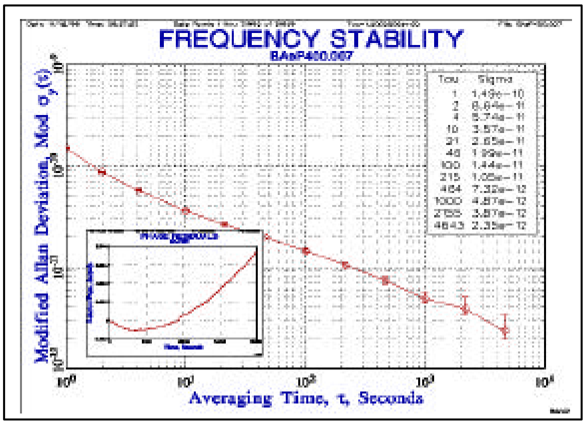

Fig 3: A Modified Allan deviation plot

TVAR, or Time Variance, was a more recent addition to the variance family and was motivated by the need for a variance concentrated primarily on time, not frequency. It was designed to look at clock performance. It is particularly suited to 'measuring the stability of time dissemination, comparison and measurements systems' **[2].** It is also ideal for measuring 'synchronisation stability of telecommunication networks' **[2]**

$$s_x^2(t) = (1/6) <(D^2 \underline{x})^2>$$

Whilst TVAR is a time stability measure, it's similar construction to MVAR ensures that it has many of the advantages of MVAR. TVAR is useful for the distinguishing of the type of noise present within time based systems, i.e. white- and flicker noise phase modulation noise.

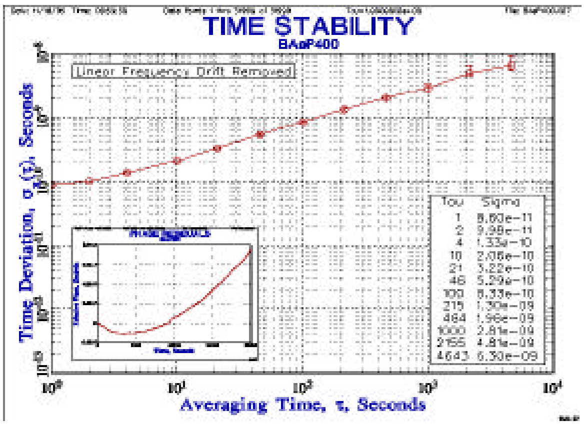

Fig 4: a Time deviation plot

Quartzlock also is able to calibrate using statistics commonly used in the telecommunications industries

Maximum TIE is calculated by moving an n-point (n= t/to) window through the phase (time error) data and finding the difference between the maximum and minimum values at each window position.

$$MTIE(t) = max[max\ x(i) - min\ x(i)],\ n=1, 2, N\text{-}1$$
$$1<k<N\text{-}n \quad k<I<k+n \quad k<I<k+n$$

MTIE is a measure of the peak time deviation of a clock and is therefore very sensitive to a single extreme value, transient or outlier.

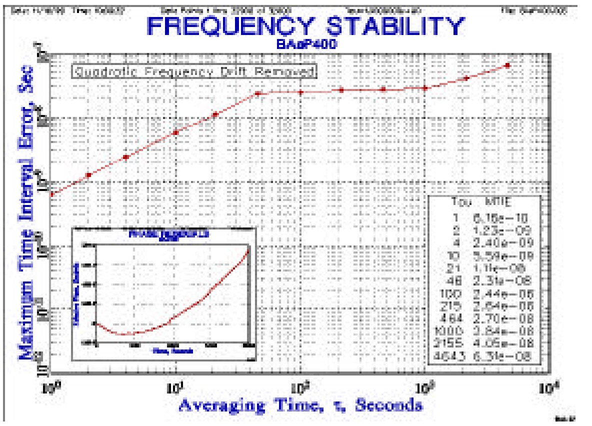

Fig 5: an MTIE plot

The **rms time interval error, TIE_{rms}** is defined by the expression:

$$TIE\ rms = SQRT\left[\frac{1}{(N-n)} \sum_{i=1}^{N-n} [x(i+n) - x(i)]^2\right]$$

For no frequency offset, TIE_{rms} is approximately equal to the standard deviation of the fractional frequency fluctuations multiplied by the averaging time. It is therefore similar in behaviour to TDEV, although the latter is better suited for divergent noise types **[13]**

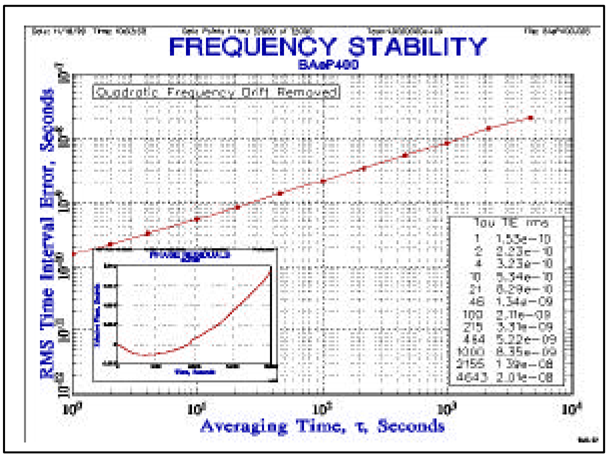

Fig 6: A rms. time interval error plot

Quartzlock use the Stable32TM data analysis program from Hamilton Technical Services. The following procedure is followed when making measurements:

i) The raw ASCII data is transferred from the floppy disk to the appropriate folder on the hard drive.

ii) The data file is opened.

iii) The raw phase data is converted from phase to frequency data.

iv) Corrections are made for scaling factors present in the measurement system.

v) The frequency data is analysed for outliers (i.e. data points a long way outside the standard deviation of the data). Quartzlock remove any data points outside 5σ, where σ is the standard deviation of the measurements. All outliers removed are accounted for by checking when the outlier occurred and reasoning why it may have happened. If there exists a significant number of outliers- say more then 10-20- then the data run is repeated. This prevents bad data being processed and conclusions erroneously drawn from it. Outliers may results from equipment failure due to mains outages, loss of lock, extreme weather conditions (especially for GPS and externally derived signals like Loran-C and Radio 4), rise in temperature etc.

vi) Both the phase and frequency data (after removal of outliers) is then plotted and visually analysed.

vii) The drift is then computed, both per day and per month.

viii) The frequency stability data is then computed for AVAR, MVAR and TVAR.

It automatically calculates τ_{min} and τ_{max} from the number of data points, calculates the noise type for each averaging factor, the confidence intervals, max, mean and min sigma values. Quartzlock generally remove the linear frequency drift, to ensure that we are characterising the short-term frequency instabilities, not long-term drift. This doesn't matter too much for Rb and H-Masers, but can significantly alter the results for an OCXO, which drifts rapidly in frequency. All results are then plotted in a standard format, as laid down in procedure for calibration at Quartzlock. This ensures results can be fairly compared between calibrations performed on different days.

ix) If requested by the customer, Quartzlock can also characterise frequency standards using more modern, yet experimental variances, like Hadamard, overlapping Hadamard, and Total Variances [14]. This is not a standard a procedure however.

x) For items destined for telecommunications applications, Quartzlock wil characterise standards using rms. and maximum time interval measures [14]. These are similar to TVAR.

xi) All results are printed out and the appropriate calibration and conformance certificates are filled.

xii) If the unit has failed an aspect of the test, then Quartzlock endeavour to investigate why this has happened and try to fix the fault. No unit is sent out without have conformed to the standards laid down by Quartzlock.

In this document, I have tried to outline and clarify some basic theory and practice of frequency (and relative time) calibration. If the user requires more detail on this matter, he is advised to refer to the documents referenced throughout the article..

.

References

[1] An Introduction to frequency calibration- Michael Lombardi, Time and Frequency Division, NIST

[2] A Frequency domain view of Time domain characterisation of clocks and time and frequency distribution systems- David W. Allan, Marc A. Weiss & James L. Jespersen, Proc. 45th Annual Frequency Control Symposium, 1991, pp 667-678

[3] Standard terminology for fundamental frequency and time metrology- David W. Allan et al, Proc.42nd Annual Frequency Control Symposium, 1988, pp 419-425

[4] Trends in International Timing- David W. Allan and Alex Lepek, Proc.7th European Frequency and Time Forum (EFTF), March 1993, pp 221-224

[5] Noise and Time And Frequency- A Potpourri- James A. Barnes, Austron Inc., Proc.42nd Annual Frequency Control Symposium, 1988. Pp 426-431

[6] Statistics of Atomic Frequency Standards, David W. Allan, Proc. Of IEEE 54 pp 221-231 (1966)

[7] Quartzlock Manual for Model A7 Frequency and Phase Comparator, April 1999, R.Percival and Dr CJM Little

[8] Quartzlock Data Sheet for A7 Frequency and Phase Comparator, April 1999, R.Percival, Dr CJM Little and CR Green

[9] Guide Technology Inc. GT200 Time Interval Counter Card manual.

[10] Time and Frequency Metrology: Current Status and Future Considerations, David W. Allan, Time and Frequency Division, NIST, Boulder, Colorado, Proc. 5th European Frequency and Time Forum (EFTF), March 1991, pp 1-9

[11] Report on the study to determine the suitability of GPS disciplined oscillators as Time and Frequency Standards traceable to the UK national time scale UTC (NPL), NPL report No CTM 1, Drs John Davis and Joan Furlong, Centre for Time Metrology, National Physical Laboratory, London, October 1997

[12] Tutorial: Clock and Clock Systems Performance Measures- David W. Allan, Allan Time

[13] Introduction to the Time Domain Characterisation of Frequency Standards- James Jespersen, Time and Frequency Division, NIST, Boulder, Colorado.

[14] Manual for Stable32, Version 1.21 by Hamilton Technical Services March, 1999