Pitfalls An Engineer Needs To Be Aware Of During Vibration Testing

Pramod Malatkar*, Shaw Fong Wong, Troy Pringle, and Wei Keat Loh
Intel Corporation
MS CH5-159, 5000 W Chandler Blvd, Chandler, AZ 85226
*Email: pramod.malatkar@intel.com, Phone: +1 480-552-8081, Fax: +1 480-554-1521

Abstract

In the electronic industry, there is currently a renewed interest in the study of the failure of BGA package and socket solder joints under mechanical stressing, specifically shock and vibration. This means that an analyst now needs to know the proper modeling and test procedures to use in order to better tackle the problem of solder joint reliability under shock and vibration loading. In this paper, an effort has been made to collate the best known methods associated with vibration testing of electronic components mounted on printed circuit boards.

Introduction

There has been a continuous reduction in the size and pitch of solder balls on components, like BGA packages, sockets, etc., that are surface mounted on printed circuit boards. For instance, the BGA pitch over generations has reduced from 1.27mm to almost 0.6mm now. On the other hand, the severity and frequency of the loading conditions that these components are subjected to has been continuously increasing, as in the event of dropping of cell phones, laptops, etc. on hard surfaces. These two factors have begun to impact adversely the solder joint reliability (SJR). In addition, the transition to lead-free solder balls further reduced the marginality of the SJR, driving OEMs to resort to corner gluing or underfill use, especially with laptop and cell phone packages. In other words, there is currently a need to study the BGA solder joint failure under shock and vibration loading. It is expected that experimentation would play a big role in helping us understand the fatigue failure of BGA solder joints. Therefore, the engineer now needs to understand the critical elements involved in shock and vibration testing.

In this paper, we talk about the best known methods associated with vibration testing of electronic components. Various aspects of vibration, including natural frequencies, mode shapes, FFT, sampling rate, effect of nonlinearities, PCB pad and trace design, etc., are touched upon to provide the analyst with the necessary information required for conducting a successful vibration test. To assist those interested in learning more about the topics included in this paper, we have also cited useful references.

Determination of Natural Frequency

The first step in the vibration testing of any system is the determination of its natural frequencies. The most common approach is to do a sine sweep at a finite rate over the interested range of frequency, usually starting at zero and stopping at several hundred Hz. In the sine sweep process, any kind of data – strain, displacement, velocity, or acceleration – could be used. In the frequency-response function of the system, the response amplitude peaks around the natural frequencies and dies down quickly away from them. Therefore, the problem of determining the natural frequencies reduces to identifying the location of these peaks in the frequency-response function. Figure 1 shows the sine sweep data collected on a board starting at 60 Hz and ending at 500 Hz and at an input level of 0.1G and a sweep rate of 30 Hz/min. However, one needs to be very careful while deciding upon the sweep direction, sweep rate, and input acceleration level. The relationship between these parameters and the measured value of the natural frequency is discussed next.

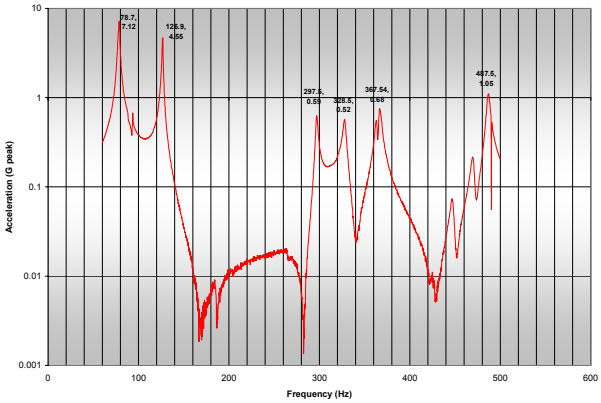

Figure 1: Typical sine sweep frequency-response function.

All systems have some degree of nonlinearity present in them. Nonlinearities could be either geometric (i.e., nonlinear restoring force), material (i.e., nonlinear stress-strain relationship), or due to damping (aerodynamic drag, Coulomb friction, etc.), inertia (centripetal and Coriolis acceleration), and discontinuities (impact, backlash, etc.) in the system. The presence of nonlinearity results in a shift in the system response away from the expected linear behavior. Typically, the deviation from linear behavior increases with an increase in the amplitude of system displacement.

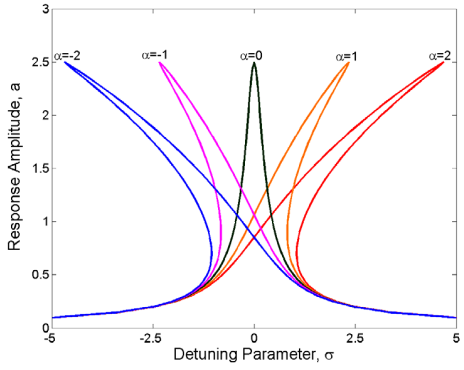

Figure 2: Effect of nonlinearity on the frequency-response curves of the Duffing oscillator.

As observed in the frequency-response curves for primary resonances of the Duffing oscillator, shown in Figure 2, nonlinearity essentially bends the frequency-response curve either to the right or to the left, depending on the type of nonlinearity. Basic characteristics of the Duffing oscillator are included in the Appendix. The degree of bending of the frequency-response curve is directly proportional to the strength of the nonlinearity. The linear behavior corresponds to the case when $\alpha = 0$.

The typical frequency-response curve of a system with *hardening* cubic nonlinearity is indicated in Figure 3. As is evident from the figure, the system now has multiple solutions over a range of frequency values which leads to a *jump* during the forward or backward sweep, as indicated by the vertical arrows. Also, depending on the sweep direction the measured maximum response amplitude would be different. In other words, different sweep directions lead to different measured values of the natural frequency.

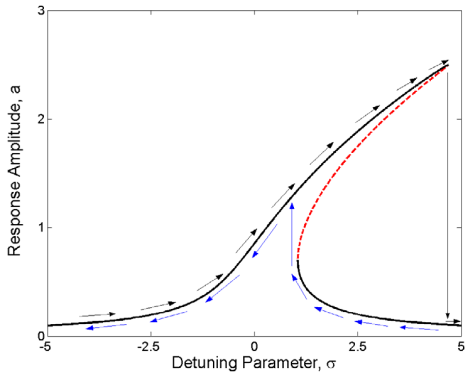

Figure 3: Frequency-response curve indicating the *jump* phenomena. Arrows indicate the sweep direction.

Another interesting thing to note in Figure 3 is that the two measured values of natural frequencies, during the forward and backward sweeps, are higher than the linear natural frequency, which corresponds to the point where the detuning parameter $\sigma = 0$. Similarly, for a system with *softening* cubic nonlinearity, the two measured values would be smaller than the linear natural frequency. And with an increase in the input G level (or forcing amplitude "f" of the Duffing oscillator), this deviation becomes even worse, as illustrated in Figure 4.

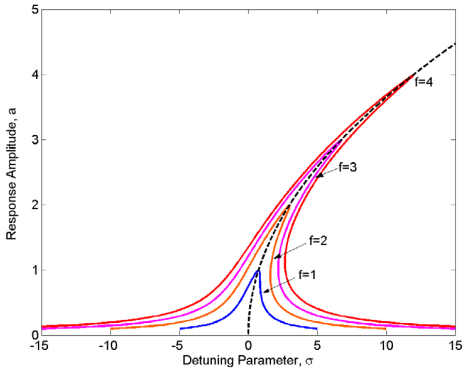

Figure 4: Effect of forcing amplitude on the system response.

In the case of a low forcing amplitude, the *jump* phenomena is absent and the system response is closer to the linear frequency-response curve, as is evident for the case f = 1 in Figure 4. In other words, to accurately determine the linear natural frequency, the input G level should be kept to a minimum.

During the sine sweep, attention should also be paid towards the sweep rate being used. High sweep rates reduce the time required for data collection but would lead to incorrect response values, as shown in Figure 5.

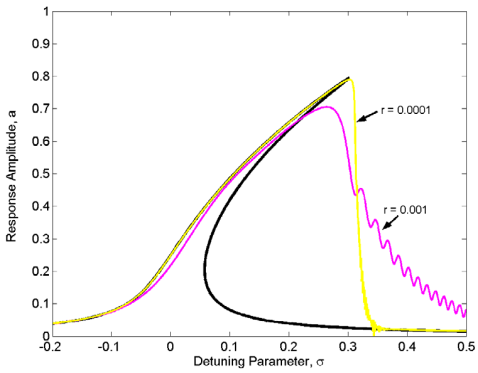

Figure 5: Effect of sweep rate on the measured system response. The thick *black* curve indicates the actual response.

Ideally, a sine sweep should be a quasi-stationary process; i.e., after every increment (or decrement) in the input frequency, a sufficient amount of time should elapse prior to noting the response, so as to allow the system to reach steady state. However, this would make the sweeping process very time consuming. Therefore, it is suggested to do two sweeps: first one at a relatively *high* sweep rate over the entire frequency range of interest, so as to get a rough idea of the natural frequencies; followed by a *low* sweep rate around the peaks observed in the first sweep, but over a smaller frequency range.

An alternate method of determining the natural frequencies is to use the so-called *bump* test. In this test, the system is hit by a hammer and an estimate of the natural frequencies is obtained from the peaks observed in the subsequent frequency spectrum. The peak-hold feature in the signal analyzer should be used to get distinct peaks in the frequency spectrum. Later, a sine sweep around these peaks at a low sweep rate would provide accurate natural frequencies.

Multiple Modes and Higher Harmonics

Any continuous system like a PCB has an infinite number of deflection modes, each with a distinct mode shape and natural frequency. Under sinusoidal or random input, the board response would typically contain significant contribution from one or more modes. In the case of random vibration, a wide range of frequencies are present in the input and that explains the presence of multiple modes in the board response. However, in the case of sinusoidal (i.e., single-frequency) input, inherent nonlinearities in the system are responsible for excitation of multiple modes. In continuous systems, nonlinearities essentially couple the linearly uncoupled modes, and this coupling leads to modal

interactions (i.e., interaction between modes), resulting in transfer of energy among modes. Modal interactions may be the result of internal resonances, combination external resonances, combination parametric resonances, or non-resonant interactions. For further details on modal interactions, please refer to [1].

More often than not, higher harmonics of the excited modes are also present in the system response. A system with cubic nonlinearity would have the odd harmonics (3, 5, 7, …) of the excited mode(s) in the response, while one with quadratic nonlinearity would have both even and odd harmonics (2, 3, 4, …) along with a DC component in the response. Figures 6 and 7 show the typical response FFT plots of systems with cubic and quadratic nonlinearity, respectively, for a sinusoidal input. It is obvious that, in addition to a large peak at the input frequency of 1 Hz, there are peaks present at the higher harmonics as well. In Figure 7, the peak at 0 Hz refers to the DC component in the response.

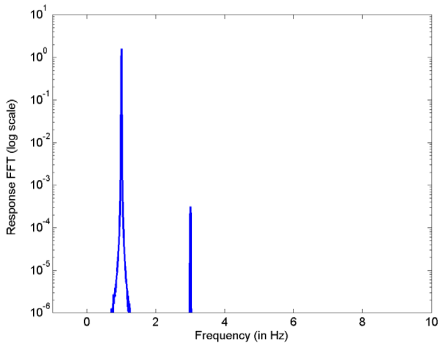

Figure 6: Response FFT of a system with cubic nonlinearity.

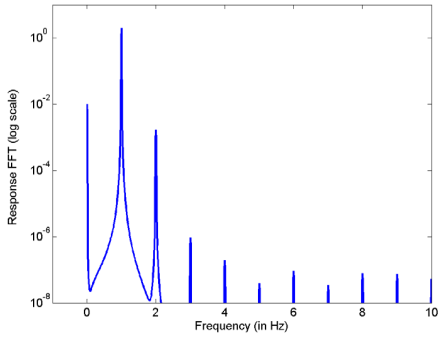

Figure 7: Response FFT of a system with quadratic nonlinearity.

Also, nonlinearity in a system could lead to a large response around the natural frequency when the input frequency is close to an integral multiple (subharmonic resonance) or sub-multiple (superharmonic resonance) of the natural frequency. For further details on nonlinear resonances, please refer to [2].

In some cases, it might be ideal to have a sinusoidal response to a sinusoidal input; i.e., not have multiple modes present in the system response which makes the response amplitude to vary with time, rather than being constant. For instance, when the ratio of the first two natural frequencies is close to 1:3 (internal resonance) and cubic nonlinearity is present in the system, then exciting one mode would indirectly excite the other. This can be avoided by slightly varying the boundary conditions (say, by changing the location of the mount points of the board to the fixture), so that the ratio of the natural frequencies is considerably away from 1:3.

Data Acquisition and Analysis

The *continuous* strain or acceleration signal is usually digitized for data storage and analysis. To be able to reconstruct the signal accurately, it is important that the signal sampling rate (i.e., number of data points stored per second) be at least twice the highest frequency component (in Hz) present in the signal, otherwise there would be *aliasing*, resulting in the sampled signal looking different than the actual signal, as illustrated in Figure 8. This requirement is known as the Nyquist criterion and the minimum sampling frequency of twice the maximum frequency in the signal is known as the Nyquist frequency [3].

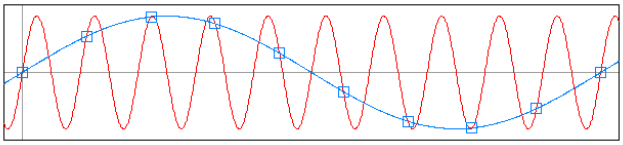

Figure 8: Example of signal aliasing. The high and low frequency waves correspond to the original and sampled signals, respectively. (Courtesy: Wikipedia [3])

To identify the modes and higher harmonics present in the system response, the commonly used tools are the Fast Fourier Transform (FFT) and the Power Spectral Density (PSD) of the response time series data. An FFT shows the spectral or frequency components present in the signal and a PSD indicates the frequency distribution of signal power [4]. The term power is used because the dynamical power in a vibrating system is proportional to the square of the vibration amplitude. Both FFT and PSD plots contain peaks at the frequencies present in the system response, which correspond to modal frequencies and their harmonics.

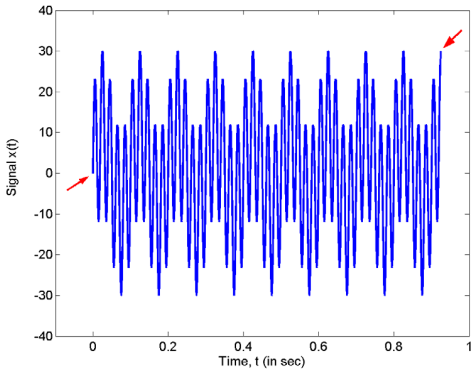

Figure 9: Example of a *non-periodic* time trace.

To get a good resolution on the frequency and amplitude of the peaks in the FFT (or PSD), windowing might be necessary. The FFT computation assumes that the time trace is periodic in the duration of the trace; i.e., the first and last points in the time series data have the same value, which mostly is not the case. In Figure 9, the time trace of a signal

comprising of only two frequency components (at 10 and 50 Hz) is shown. In this time trace, the first and last data points (indicated by arrows) have unequal values, and hence it is considered to be *non-periodic*.

When the FFT of a non-periodic time trace is computed, then the resulting frequency spectrum suffers from leakage. Leakage results in the energy smearing out over a wide frequency range in the FFT when it should be in a narrow frequency range. Since most time traces turn out to be non-periodic, a *window* must be applied to correct for leakage. A *window* is essentially a weighting function that makes the start and end points equal to zero; thus rendering the *modified* time trace periodic. In Figures 10(a) and (b), FFT plots obtained with and without a window are illustrated. The non-periodic time trace shown in Figure 9 was used to generate these plots. Figure 10(a), obtained using the Hanning window, shows the two frequency components present in the signal as distinct peaks in the FFT plot. On the other hand, Figure 10(b) indicates that the energy is distributed over a wide frequency range, which is incorrect.

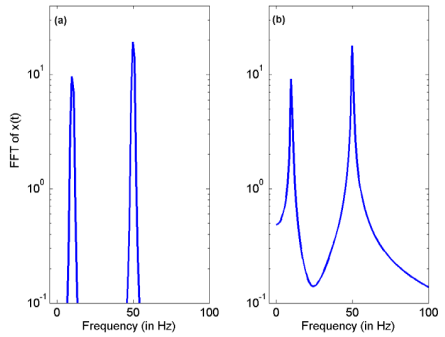

Figure 10: FFT of a *non-periodic* time trace obtained (a) using the Hanning window and (b) without use of any window.

It should be noted that windowing is a post-data collection step, but done prior to FFT (or PSD) computation. Most signal analyzers have windowing capability and also software tools like MATLAB provide window functions. For an accurate frequency measurement a Hanning window is recommended, where as for precise amplitude measurement a flat-top window is preferred. For more details, the reader is referred to [5].

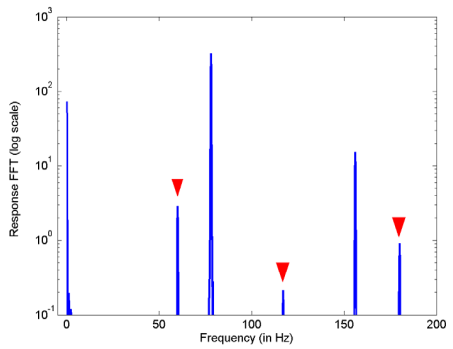

Figure 11: FFT indicating electrical noise components.

In addition to peaks at the modal frequencies of the excited modes and their higher harmonics, peaks can also be present at 50 Hz, 100 Hz, 150 Hz, etc. in the FFT plot. These *extra* peaks correspond to the fundamental and higher harmonics of the electrical noise present in the signal. In some countries, the electrical noise peaks appear at 60 Hz, 120 Hz, 180 Hz, etc. Figure 11 illustrates the noise present in the FFT plot of a board strain data collected in Singapore. The noise related peaks at 60 Hz, 120 Hz, and 180 Hz are indicated by the red triangles. Multiple grounds at different potentials (or presence of *ground loops*) is the primary cause for this type of noise and the best way of avoiding this problem is to have a single ground. For more information on *ground loops* and the different types of noise signals in general, the reader is directed to [6].

All the discussion in this section so far basically indicated that in an FFT (or PSD) plot of a response time trace, there are usually multiple peaks present at different frequencies. Each peak does not necessarily correspond to a distinct mode; it could correspond to a higher harmonic of the excited mode(s) and/or electrical noise. Therefore, care should be taken while deducing the modes present in the system response based on the FFT. Another note of caution is to never apply an FFT on minimum or maximum principal strain data, as these do not correspond to strain at a fixed location and orientation in the system and therefore do not represent the system dynamics.

In addition to knowing the natural frequencies of modes, having an idea of their corresponding deflection shapes is also useful. A convenient way of determining the mode shapes is through the strobe light test. In this test, the system is excited close to the natural frequency of the mode of interest and the frequency of the strobe light is matched with the excitation frequency of the shaker. In such a case, the system deformation surface would appear frozen in time and would be a close approximation of the mode shape. The strobe light test, however, works well only on smaller structures and that too in darkness. Also, the mode shape information cannot be stored for future analysis. A more advanced method is to use a high-speed camera, which however is expensive. An alternate approach to obtain the mode shapes as well as natural frequencies is through FE modeling of the system.

In some systems, there could be a continuous static load present on the board, due to the assembly fixture or heatsink. Any kind of pre-load on the board would result in a DC component in the strain data. In the time-series data, this would result in a non-zero mean and, it would appear as a peak at zero frequency in the FFT plot.

PCB Design

To study the solder joint reliability of BGA package (or socket) surface mounted onto a PCB, a daisy chain structure in the package and board is commonly used. During the vibration test, this daisy chain is monitored in real time for an electrical open which corresponds to failure of solder joint(s). But it has been observed that, in addition to solder joint failures, it is also possible for the pads and/or traces in the PCB to fail under vibration loading. Figures 12 and 13 are

cross-section pictures indicating pad crater and trace cracking, respectively.

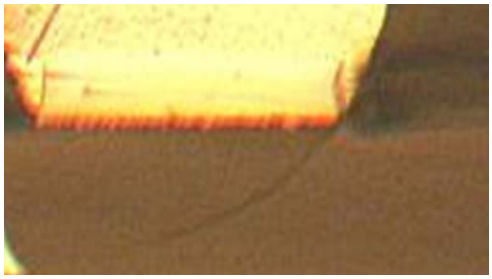

Figure 12: Partial pad crater on PCB side.

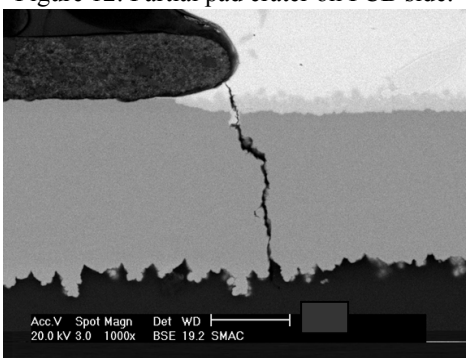

Figure 13: SEM image indicating 100% trace crack.

Like most solder joint failures, failures due to pad crater and trace cracking also occur mostly in the corner areas of the BGA package. Fortunately, by some simple trace and/or pad design changes in the high-risk areas, these failures can be avoided. A *snow-ball* (or *modified tear-drop*) pad design, shown in Figure 14, significantly reduces the risk of trace cracking. However, a simpler alternative (from board layout design perspective) is to use a *thick* trace, which is shown in Figure 15. The recommendation is to use a trace width as close to the pad diameter as possible for a minimum of 6 mils extension from the edge of the BGA pad to the end of the trace segment. It has also been observed that the orientation of the trace plays a role in trace cracking. Depending on the shape of the dominant deflection mode, an optimized trace orientation can be easily determined. Moreover, using solder mask-defined pads, instead of metal-defined pads, would help reduce the risk of failure due to trace crack and to some extent pad crater also. Another ingenious way of circumventing the problem of trace cracking is to use a *redundant* trace wherever possible; thus, if one trace fails the other would still keep the electrical connectivity.

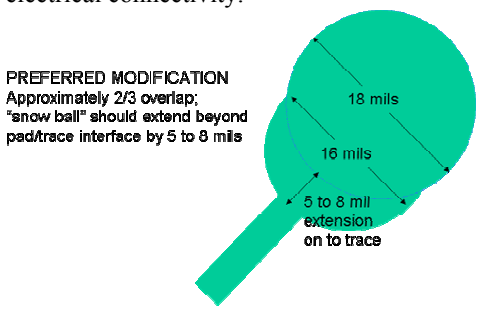

Figure 14: Snow-ball pad design.

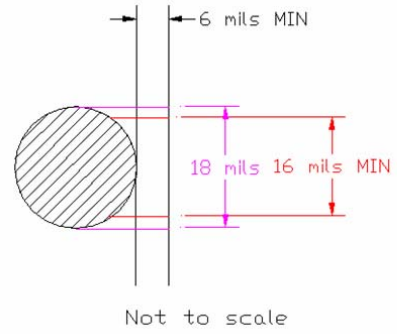

Figure 15: Recommended *thick* trace design.

It is well known that most of the solder joint failures are due to tensile loading of the solder joints, which however is very hard to measure. It has been found that the compressive strain on the bottom side of the board gives a good indication of the severity of the tensile loading of the solder joint on the other side of the board. Therefore, board strain monitoring via strain gages glued onto the bottom side of board is becoming increasingly popular in shock and vibration tests [7].

Through modeling and experimentation, it has been observed that the board strain varies significantly in the BGA package corner areas, where most of the solder joint failures occur. This is illustrated in the strain distribution plot in Figure 16, which was generated using FE modeling. In other words, a 1mm offset in the location of the strain gage could result in a difference of several hundred micro-strain in the measured strain response. Therefore, it is very important to ensure that the strain gage is attached at a consistent location on the PCB, so as to allow comparison of results between different boards. For this purpose, it is suggested to print a silk screen on the PCB surface to aid in proper placement of the strain gage.

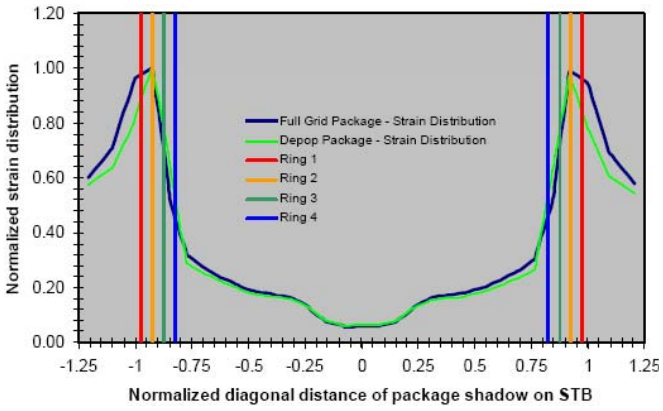

Figure 16: Normalized board strain distribution in and around BGA package area.

It should be noted that no two boards are identical and therefore a variation in the natural frequency between boards of up to 5% is typical. Even such a small variation could result in significantly different strain responses for the same input G level and input frequency near resonance. Therefore, in such cases, attaching strain gages onto each of the boards is

recommended, especially in tests where strain values between boards would be compared.

Finally, we would like to provide some means of reducing the solder joint related failures in boards during shock and vibration testing. From beam and plate theories, we know that the board strain is proportional to its curvature at that point. Therefore, to reduce the strain value (and thus the risk of solder joint failure), the board displacement amplitude should be kept small. This can be achieved in one of the following ways: (a) by increasing the board thickness; (b) by adding metal ribs or stiffeners if there is room to do so; (c) by adding intermediate support points in addition to the support provided at the edges; and (d) by adding active and/or passive isolators so as to reduce the amount of energy transferred from the source to the board. For further details on this topic, the reader is referred to [8].

Conclusions

In this paper, an attempt has been made to intimate engineers of pitfalls involved in the vibration testing of electronic components. Most of these recommendations would be applicable to shock testing as well. Following the procedures described here would aid in: (a) the design of PCB pad and traces that perform better in vibration tests; (b) accurate data collection; and (c) better data and failure analysis. In other words, it would help eliminate the need for redesign of PCB and/or retesting; thus, resulting in a drastic cut in the time, effort, and cost incurred in the testing of electronic components under vibration loading.

Acknowledgments

We would like to acknowledge the technical support and encouragement received from Ward Scott and Luke Garner. Inputs received from Karumbu Meyyappan, Rene Sanchez, Chad Phipps, Hong Keat Leow, Tozer Bandorawalla, Rick Canham, and Shankar Ganapathysubramanian, especially on PCB trace and pad designs, are also acknowledged.

References

1. Nayfeh, A. H., Nonlinear Interactions, Wiley (New York, 2000).
2. Nayfeh, A. H. and Mook, D. T., Nonlinear Oscillations, Wiley (New York, 1979).
3. Wikipedia, http://www.wikipedia.org.
4. Harris, C. M. and Crede, C. E., Shock and Vibration Handbook Volume 1 – Basic Theory and Measurements, McGraw-Hill (New York, 1961).
5. LDS, Application Note AN014 on Understanding FFT Windows, http://www.lds-group.com/docs/site_documents/AN014 Understanding FFT Windows.pdf
6. CAPGO Pty Ltd, Document on Understanding Noise in Measurement, http://www.capgo.com/Resources/ Measurement/Noise/Noise.html
7. Liang, F. Z., Williams, R. L., and Hsieh, G., "Board strain states method and FCBGA mechanical shock analysis," *Proc 1st International Conference and Exhibition on Device Packaging*, Scottsdale, AZ, March 2005.
8. Steinberg, D. S., Vibration Analysis for Electronic Equipment, Wiley (New York, 2000).

Appendix

The Duffing oscillator is a single-degree-of-freedom system with cubic nonlinearity. Assuming the damping, forcing and nonlinearity are "weak", its equation of motion is given by:

$$\ddot{x} + \omega^2 x + 2\varepsilon\mu\dot{x} + \varepsilon\alpha x^3 = \varepsilon f \cos(\Omega t),$$

where $x(t)$ denotes the system displacement, ω is the natural frequency, μ is the damping coefficient, α denotes the nonlinearity coefficient, f denotes the forcing amplitude, Ω is the forcing frequency, t denotes time and ε is a small positive parameter ($\varepsilon \ll 1$).

When $\alpha > 0$, we have a "hard" spring (or *hardening* cubic nonlinearity), and $\alpha < 0$ means a "soft" spring (or *softening* cubic nonlinearity). At primary resonance, the forcing frequency is very close to the natural frequency, and the closeness is indicated by a detuning parameter σ defined as follows: $\Omega = \omega + \varepsilon\sigma$. In this case, the system response is given by:

$$x(t) = a\cos(\Omega t - \gamma) + O(\varepsilon),$$

where a and γ represent the response amplitude and phase, respectively. All the plots related to the Duffing oscillator in this paper are for the primary resonance case only.

To learn more about the Duffing oscillator, please refer to [2].