doi: 10.1111/j.1460-2695.2006.01065.x

Review of methods to predict solder joint reliability under thermo-mechanical cycling

S. RIDOUT* **and C. BAILEY**

School of Computing and Mathematical Sciences, University of Greenwich, Park Row, Greenwich, London, SE10 9LS, UK

Received in final form 20 July 2006

ABSTRACT Solder joints are often the cause of failure in electronic devices, failing due to cyclic creep induced ductile fatigue. This paper will review the modelling methods available to predict the lifetime of SnPb and SnAgCu solder joints under thermo-mechanical cycling conditions such as power cycling, accelerated thermal cycling and isothermal testing, the methods do not apply to other damage mechanisms such as vibration or drop-testing. Analytical methods such as recommended by the IPC are covered, which are simple to use but limited in capability. Finite element modelling methods are reviewed, along with the necessary constitutive laws and fatigue laws for solder, these offer the most accurate predictions at the current time. Research on state-of-the-art damage mechanics methods is also presented, although these have not undergone enough experimental validation to be recommended at present.

Keywords constitutive law; fatigue; finite element; life prediction; solder joint; thermal cycling.

NOMENCLATURE

A = crack area
a = joint diameter
$\frac{da}{dN}$ = crack propagation rate
c_i = material constants (Unless stated otherwise, similarly numbered constants in different equations are not the same, e.g. c_2 in Eq. (2) is not equal to c_2 in Eq. (13))
D = damage
E = Young's modulus (elastic modulus)
g = cohesive zone stiffness
h = solder joint height
L = distance to neutral point
N_f = mean number of cycles to failure
N_{ff} = number of failure free cycles
N_0 = number of cycles to crack initiation
Q = activation energy
R = universal gas constant
T = temperature
ΔT = difference in temperature between the hot and cold extremes during cycling
ΔW = accumulated strain energy density per cycle
Y = cohesive zone traction
α = CTE (coefficient of thermal expansion)
$\Delta\alpha$ = difference in CTE between the circuit board and the component

* *Correspondence*: S. Ridout. E-mail: s.w.ridout@gre.ac.uk

 © 2006 The Authors. Journal Compilation © 2007 Blackwell Publishing Ltd.

$\Delta\gamma$ = shear strain range
ε = normal strain
$\varepsilon_{\mathrm{acc}}$ = accumulated effective plastic strain per cycle
$\varepsilon_{\mathrm{cr}}$ = total creep strain
$\varepsilon_{\mathrm{in}}$ = instantaneous plastic strain
ε_{p} = total plastic strain
ε_{s} = steady-state creep strain
$\varepsilon_{\mathrm{th}}$ = thermal strain
η = characteristic life
μ = cohesive zone relative opening displacement
σ = normal stress

INTRODUCTION

Most electronic devices are expected to last for many years of use. Therefore it is unfeasible for companies to perform field tests due to the time and cost involved. Instead, accelerated tests are used which impose much harsher conditions on the joint, causing it to fail in a shorter time and allowing a judgement to be made on the products reliability. The most popular accelerated test is thermal cycling, in which a soldered assembly is placed in an oven and subjected to a cyclic temperature profile representing conditions considerably harsher than expected field use conditions. During cycling or at periodic intervals the samples will be monitored for electrical resistance/continuity, if the resistance shows a very large increase this indicates that a crack has grown completely through the solder joint and the sample is considered to have failed. The average number of cycles to failure in such a test is used as a measure of the reliability of a chip assembly.

Another accelerated test is the power cycling test, during which heat is periodically generated from within a component. This results in an anisothermal temperature field more closely matching the conditions experienced in actual field use.

Two of the more common types of chip assemblies tested are surface mount passives and flip-chips shown in Fig. 1. The circuit board is typically FR4 but alumina or other materials may be used. Two of the more common solder alloys are SnPb, which due to its Pb content is banned from use in most applications throughout Europe from July 2006 under the RoHS directive,[1] and SnAgCu which is the most popular Pb-free replacement. Throughout the paper SnPb refers to eutectic Sn37Pb and SnAgCu refers to Sn3.5Ag0.7Cu or a slight variation on this composition. Although many of the methods reported in this paper were developed for a particular solder alloy, in most cases the same methods could be used for a different alloy if the material constants are changed.

The primary failure mechanism of solder joints is ductile fatigue fracture. During operation or during testing, an electronic component will be subject to changes in temperature. This causes the materials to expand and contract at different rates depending on their coefficients of thermal expansion. This cycling results in internal stresses developing which, in turn, causes the solder to creep. Creep is a time-dependent plasticity, which occurs in metals when they are close to their melting points. Over many cycles, the creep leads to ductile fatigue damage, which manifests itself as microvoids in the solder, which grow and coalesce into macro-cracks that slowly propagate through the joint over the course of its lifetime (Fig. 2), ultimately resulting in failure of the joint.

It should be noted that other failure mechanisms are possible depending on the external conditions applied. For instance, if the device was subject to vibration or dropped onto the floor then the sudden acceleration may cause a brittle failure.

MODELLING METHODS

Modelling is a useful tool used to supplement or replace accelerated tests, particularly in the early design stages. The modelling discussed in this paper applies to creep induced ductile fracture only, and so will be suitable for modelling the damage that occurs due to typical thermal or power cycling of joints, but not to vibration or drop testing, or shear-strength testing. Most of the methods described in this paper can be applied to any solder alloy provided the correct material constants are known. Many methods exist, and there is no clear-cut answer as to which is best. Figure 3 shows the classes of modelling methods which will be discussed in this paper.

The methods in Fig. 3 increase in complexity going from top to bottom, the *analytical* method proposed by Engelmaier[2,3] being simple to implement but with many caveats restricting its use in certain situations. The *constitutive law + fatigue law* class of methods (encompassing FEA and other alternatives) are very popular, providing more accurate predictions with fewer restrictions than *analytical* methods, however, with increased set up time and computational cost. The *damage mechanics* based methods require considerably more effort both in implementation and computational cost and their predictive capability is

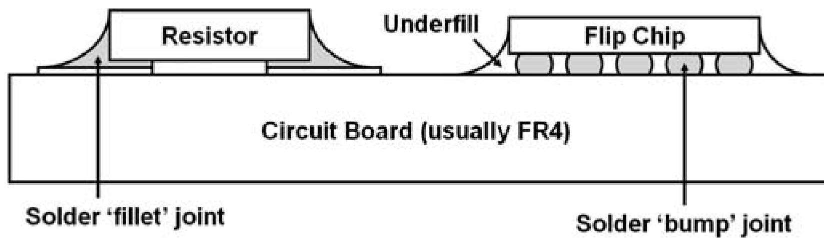

Fig. 1 Solder joints in surface mount resistor and flip chip packages.

currently unproven, however, they promise to provide the most accurate predictions once they are developed further.

In reality the behaviour of solder joints is very complex and all the modelling methods discussed must make many simplifications and omit certain aspects of the problem, for example the behaviour of the intermetallic particles in SnAgCu, or the differences between the Pb-rich and Sn-rich phases of SnPb. The *constitutive law + fatigue law methods* make fewer simplifications than the *analytical methods*, and the *damage mechanics* methods make fewer simplifications still but they all make simplifications. Many phenomena will not be explicitly modelled, and any changes in the effect of these phenomena when using a different geometry, temperature range or other condition will not be captured. (e.g. The intermetallic layer is usually not explicitly modelled, this is fine for most conditions but will yield inaccurate predictions if modelling extremely thin joints where the intermetallic layer forms the bulk of the joint.)

The methods discussed are all intended to predict the results of accelerated tests and not field-use reliability. They cannot accurately predict field-use reliability as they have not been validated against field-use data, because generally field-use data do not exist. The assumption is that the relative reliability of various components under accelerated conditions will be the same, or at least similar, under field-use conditions. The models have been extrapolated to simulate field-use conditions,[4] but this is potentially misleading.

ANALYTICAL METHODS

Clech[5] has reported the use of an analytical method where the strain range in the solder is calculated assuming the solder joint is totally compliant:

$$\Delta\gamma = \frac{L \Delta\alpha \Delta T}{h}. \tag{1}$$

This strain range is then used to predict the characteristic lifetime per unit crack area of a single joint using a Coffin–Manson fatigue law with an added crack area adjustment:

$$\frac{N_f}{A} = c_1 (\Delta\gamma)^{-c_2}. \tag{2}$$

Predictions made using this method were correlated with a data set of 27 experimental data points covering different kinds of assemblies under different thermal profiles for SnAgCu solder. Despite not taking into account the creep behaviour of the solder at all, the accuracy of the predictions was in the range of $\pm 2\times$.

A slightly different analytical approach is Engelmaier's model for predicting the lifetime of Sn37Pb joints recommended in the IPC-D-279 standard.[3] There are versions for both leaded and non-leaded components (*leaded* as in gull-wing *lead*, not to be confused with Pb) and they are very simple to use when compared with FEA modelling.

The way it works is to assume the solder deforms to its maximum amount based on the CTE (coefficient of thermal expansion) of the substrate and component. This strain range is calculated for non-leaded joints using Eq. (3), this is very similar to the Eq. (1) used by Clech but with an added empirical factor.

$$\Delta\gamma = c_1 \frac{L}{h} \Delta(\alpha \Delta T). \tag{3}$$

The strain range $\Delta\gamma$, calculated using Eq. (3) is not accurate as it does not take into account the stiffness and creep properties of the solder. The solder stiffness would prevent the strain range from ever reaching this value. To compensate for this, the standard Coffin–Manson fatigue law for predicting N_f (number of cycles to fail) has been modified to include temperature and frequency effects:

$$N_f = \frac{1}{2} \left[\frac{\Delta\gamma}{2c_1} \right]^{1/c} \tag{4}$$

$$c(T, f) = -0.442 - 6 \times 10^{-4} \bar{T}_{SJ} + 1.74 \times 10^{-2} \ln\left(1 + \frac{360}{t_D}\right). \tag{5}$$

Where $\bar{T}_{SJ}$ is the mean cyclic joint temperature for which a formula is provided.[3] This approach has been shown to predict lifetimes with an accuracy of $\pm 2\times$ under the appropriate conditions.[2] However, there are many conditions in which this law does not apply. A list of caveats from the IPC standard[3] are stated below along with a brief discussion of whether they also apply to the *constitutive law + fatigue law* methods.

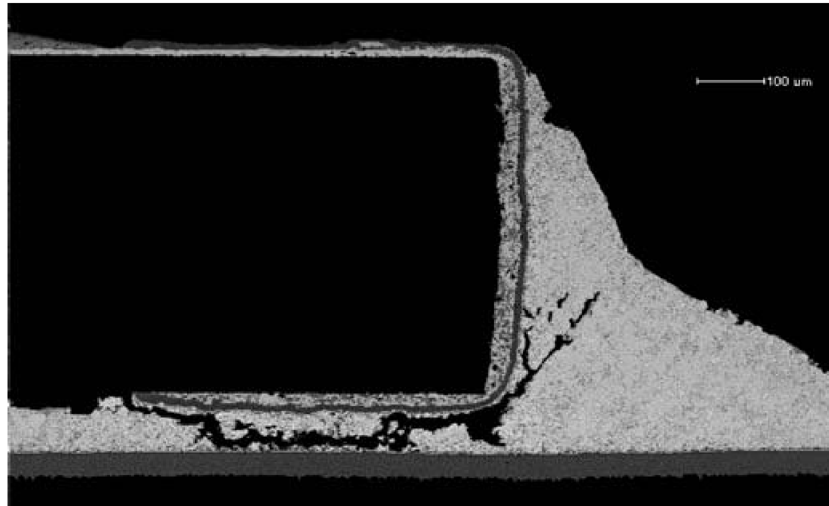

Fig. 2 SEM image of a crack through a SnAgCu solder joint.

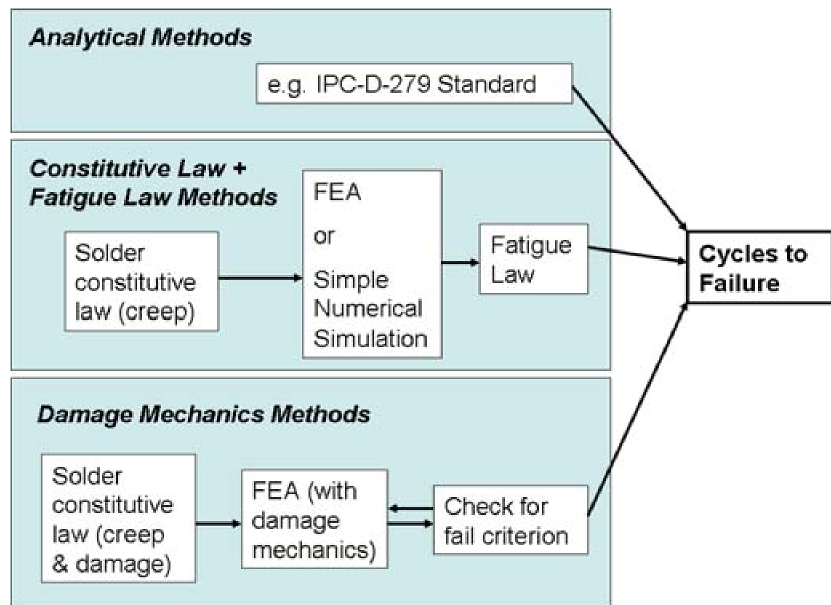

Fig. 3 Modelling methods to predict cycles to fail in thermo-mechanical cycling.

1 Solder joint quality: This caveat is to ensure that the failure mechanism is actually ductile fatigue within the solder joint and not brittle failure at the interface which could be caused by poor materials choice (e.g. alloy 42 leads) or very small solder joint gaps (<75 nm). This caveat could be applied all the modelling methods covered in this paper, which are intended to predict the lifetime of joints which fail due to creep induced ductile fracture, not brittle failure.

2 Large temperature excursions: The damage mechanism changes in solder joints experiencing large temperature excursions (-50 °C to $+80$ °C). Using a *constitutive law + fatigue law* method could capture the difference in solder behaviour within this temperature range. For example, Darveaux *et al.*[4] validated a fatigue law for SnPb for temperature ranges up to -55 °C to $+125$ °C, and Syed[6] validated a similar law for SnAgCu up to the same range.

3 High frequency/Low temperature: For frequencies >0.5Hz and/or temperatures <0 °C the direct application of Coffin–Manson may be more appropriate ($c = \sim -0.6$). Under these conditions, creep induced fatigue may not be the primary failure mechanism, in which case none of the modelling methods discussed in this paper would be suitable.

4 Local expansion mismatch. The local CTE mismatch between the solder joint and the board or component is not taken into account. For instance, an alumina component on an alumina substrate would generate no global strain in the solder and would, according to this model, experience no fatigue. However in reality, local strains will be generated

at the solder/component and solder/substrate interfaces, which would lead to damage. This damage mechanism is captured perfectly well by using FEA modelling with an appropriate constitutive law and fatigue model.

5 Very stiff leads/very large expansion mismatches: The analytical method does not make use of the Young's modulus of any of the materials so it cannot predict accurately the amount of strain which will occur in the joints, leads, component and substrate. When the package geometries and materials are similar this may not matter, but deviations from the norm could result in very inaccurate results. For instance, if we consider two assemblies, similar in all respects except that the component Young's modulus is high in one case and low in the other. The component with the high Young's modulus will cause a greater strain in the solder joints, but this model will predict the same lifetime for both. Using FEA will predict the strains occurring in the solder accurately based on the Young's modulus and other mechanical properties of the materials.

So of the five caveats, three would no longer be necessary if a *Constitutive law* + *Fatigue law* method were used instead. The *Analytical* methods, while simple to use, are limited in their applicability compared to the other methods.

CONSTITUTIVE LAW PLUS FATIGUE LAW METHODS

These methods work by running a transient simulation to predict the solder's stress strain behaviour during a thermal cycle. From this simulation, either the accumulated effective plastic strain per cycle (ε^{acc}) or the accumulated strain energy density per cycle (ΔW) is extracted to be used in a fatigue law. First, the methods of modelling the mechanical response of an assembly under thermal cycling will be discussed (FEA and alternatives), followed by the various constitutive laws available, and finally the fatigue laws.

5. FEA

FEA (finite element analysis) is a powerful and widely used numerical method, which can be used to accurately predict the mechanical response of the solder joints and surrounding assembly under thermo-mechanical cycling. A detailed description of the FEA method is provided by Zienkiewicz.[7]

To use FEA requires considerably more investment in time than the analytical approaches mentioned so far, and detailed knowledge of the test being modelled. In order to use FEA to determine the fatigue life of solder joints under thermal cycling, the following steps are required:

1 **Create geometry**
The geometry of the model should strike an acceptable balance between accuracy and simplicity. Common simplifications to the geometry include:
- The intermetallic layer formed between the solder joint and the copper pads is usually ignored, due partly to its size and partly to the lack of material property data.
- Good judgement is needed in deciding what details of the actual geometry are important to include in the model and which are superfluous.
- Symmetry. Usually only a quarter of a surface mount resistor needs to be modelled, and only an eighth of a BGA or flip-chip.

2 **Create mesh**
It is important that the mesh is fine enough to make the solution mesh independent. This can be verified by performing simulations using progressively finer meshes until the results no longer change. It is also important that the aspect ratio of elements is not too high as this can introduce numerical errors.

3 **Specify material properties**
It is crucial to use accurate material properties to obtain accurate results. When available, temperature-dependent material properties should be used. The solder properties are discussed later in the paper, because the focus is on solder joint modelling the properties of the other materials involved are not discussed.

4 **Specify boundary conditions**
During typical thermal cycling regimes the temperature may be regarded as uniform throughout the sample. For very rapid thermal shock, or when there is heat generation from within a chip (e.g. during power cycling) then the temperature field will not be uniform and should be calculated throughout the mesh on each time step.

5 **Apply fatigue law**
To apply the fatigue law, a ε_{acc} or ΔW value needs to be extracted from the FEA simulation results. Because the variation in strain across a joint is captured, this raises the question of whether to average over the whole joint or only a portion of it. Two sources have been found which address this,[6,8] and both use an average value over a 25 μm layer at the top of the solder bump in a BGA package. Altering the size of this volume averaged region may alter the value of ε_{acc} or ΔW significantly, so if the size needs to be changed, for example when modelling a different sized or shaped joint, then the fatigue law constants should ideally be re-calibrated with experimental data to ensure accurate predictions.

Alternative numerical methods

Alternatives to FEA are usually more simplified methods that capture the creep behaviour of the solder using a constitutive law but do not require as much work or

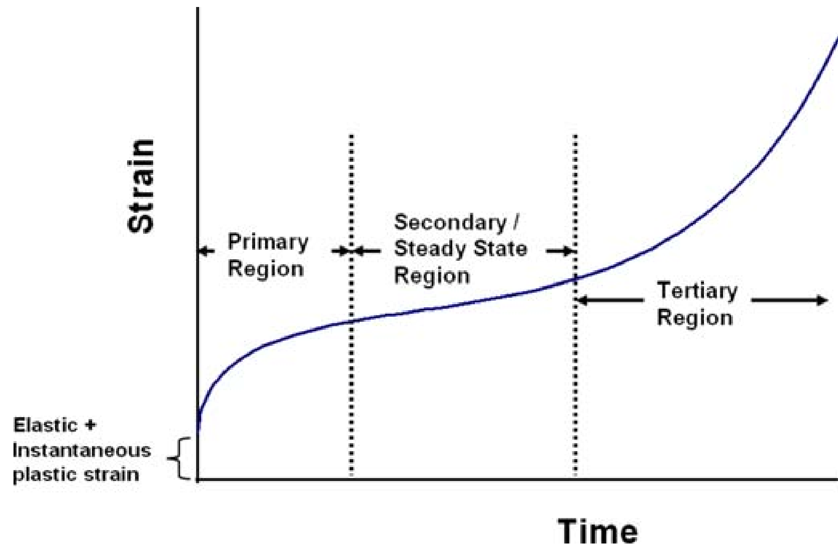

Fig. 4 Diagram showing the response of solder during a creep test.

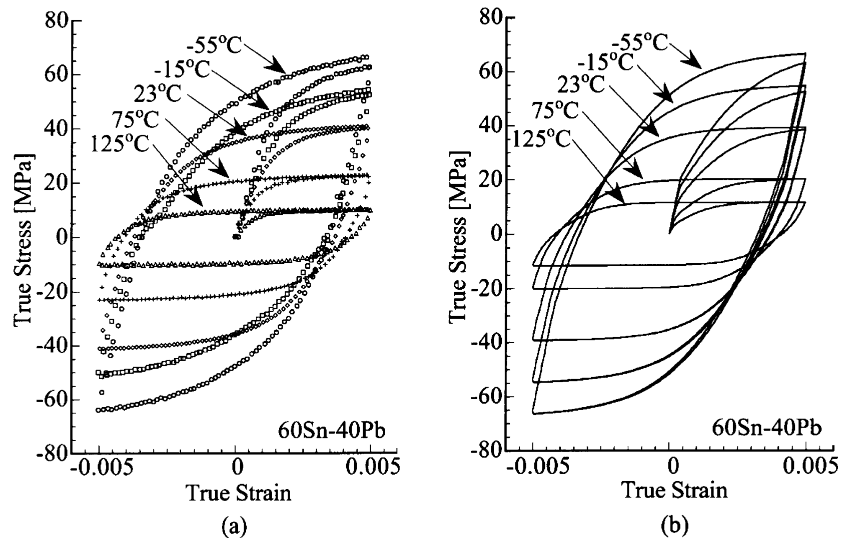

Fig. 5 Experimental and simulated hysteresis curves using the McDowell UCP model.

expertise to set up. The simplest example of this kind of approach would model only the prominent shear component of strain in the solder joint, this is an approximation as in reality a BGA solder bump would experience significant normal and shear strains. Furthermore, the strain is assumed to be uniform throughout the solder joint thus making it suitable to model global strains only.

A simulation is run where the temperature cycles between two extremes and at each time step the stress and creep strain rate in the solder are calculated. After two or three thermal cycles the response of the solder will have stabilized and the simulation can finish. The output of the simulation is a hysteresis loop – a graph of stress versus strain, and values of ε_{acc} and ΔW. Due to its simplicity it will provide results with less effort and less computation time than FEA, and because it uses a constitutive law to model the solder's creep behaviour it should provide more accurate results over a wider range of conditions than an analytical method, although no validation against experiment has been found to back this up.

A more advanced example of this kind of approach is provided by Clech[9] which provides the capability of modelling local as well as global thermal expansion mismatches, and the effect of underfill. It has been calibrated against 19 experimental results with an error of about $\pm 2.5\times$, and validation was performed against a further 14 experimental results.

A hybrid method is suggested by Darveaux[8] where two purely elastic analyses are conducted to provide (1) the stiffness of the assembly surrounding the joint and (2) the displacement between the top and bottom pads assuming the solder joint is not present. These are then used in a 1D numerical model, which uses a constitutive law to predict the average response of the solder joint over a thermal

cycle. This strikes a compromise between the accuracy of FEA and the computational ease of a simpler numerical method. However, it will not predict the damage caused by the local thermal expansion mismatch between the solder and the component or substrate.

5. Solder constitutive laws

A constitutive law for a material will govern its mechanical behaviour when subject to different stresses at different temperatures. The importance of using the correct constants for the constitutive laws used in modelling is highlighted by a sensitivity study[10] which shows that the differences in SnPb Young's modulus, CTE and creep activation energies within the range of variation reported in the literature can lead to inaccurate lifetime predictions when used in FEA modelling.

Elasticity

The elastic properties of solder are usually determined from a strain rate controlled tensile test. From this test the Young's modulus can be determined along with the ultimate tensile/shear strength of the solder. There is a large strain rate and temperature effect and the Young's modulus will appear lower at lower strain rates or higher temperatures.[11] The temperature dependence is usually modelled but no work was found which modelled the strain rate dependence.

The elastic deformation of the solder is governed by the Young's modulus and Poisson's ratio. In one dimension the relationship between stress and strain is given by Hooke's law:

$$\sigma = E\varepsilon. \tag{6}$$

In 3D the relationship is more complex[12] and takes into account the Possion's ratio (the proportion by which the cross-section area shrinks for each unit strain applied). The Young's modulus is both temperature and strain rate dependent, and a formula for calculating the Young's modulus of SnAgCu provided by Pang:[13]

$$E(T, \dot{\varepsilon})_{\mathrm{SnAgCu}} = (-0.0005T + 6.4625)\log\dot{\varepsilon} + (-0.2512T + 71.123). \tag{7}$$

The temperature dependence is usually modelled but no modelling work was found which included the strain rate dependence. More temperature dependent Young's modulus values of SnAgCu have been reviewed by Syed,[6] there is some scatter in the values reported which may be as a result of differences in the strain rate used. It is pointed out by Basaran and Jiang[14] that when modelling SnPb, the use of different values of Young's modulus (literature values range from 9GPa to 48GPa) can adversely affect the results of a simulation and that ideally a modulus measurement should be made on an actual manufactured package using a nano-indentation technique.

The CTE of the solder governs the amount by which it expands under changing temperatures:

$$\alpha = \frac{d\varepsilon_{\mathrm{th}}}{dT}. \tag{8}$$

A CTE of 25 ppm/°C has been reported for SnPb,[15] and a CTE of 20 ppm/°C has been reported for SnAgCu.[16]

Creep

Creep is a time-dependent plastic deformation, which occurs to metals under stress at high homologous temperatures, because solder has a melting point of 183 °C (SnPb) or 217 °C (SnAgCu) then it creeps at room temperature. During a creep test a constant load is applied to a solder specimen at a constant temperature. The character of the response of the solder is illustrated in Fig. 4. At the instant the force is applied, the solder will experience a strain, which is part elastic and part plastic. The elastic part can be predicted given the Young's Modulus of the solder as discussed earlier. The plastic part is the instantaneous plasticity, Darveaux *et al*.[4] and Wiese and Rzepka[17] offer laws to predict this, although the distinction between this instantaneous plasticity and the following primary creep region is not well defined. As the test continues the strain increases, first rapidly (primary region) and gradually slowing to a steady strain rate (secondary region). This additional strain is due to creep. After the secondary region comes the tertiary region during which the strain rate increases until rupture. This occurs due to both necking (shrinking of the cross section resulting in increased stress) and damage (cracking) occurring in the solder.

The standard method for FEA modelling of solder joints involves modelling only the steady-state creep and it is possible to get reasonable predictions using this approach.[4,6] However, it is possible that better accuracy may be achieved by incorporating primary creep and instantaneous plasticity into the model. (If the goal of a simulation is to accurately predict the amount of deformation occurring in the solder, then including primary creep is essential.)

Steady-state creep

The secondary creep, also called steady-state creep, is commonly the only kind of creep to be modelled. The simplest steady-state creep law is the Norton law:

$$\dot{\varepsilon}_{\mathrm{cr}} = c_1\sigma^{c_2}\exp\left(\frac{-Q}{kT}\right). \tag{9}$$

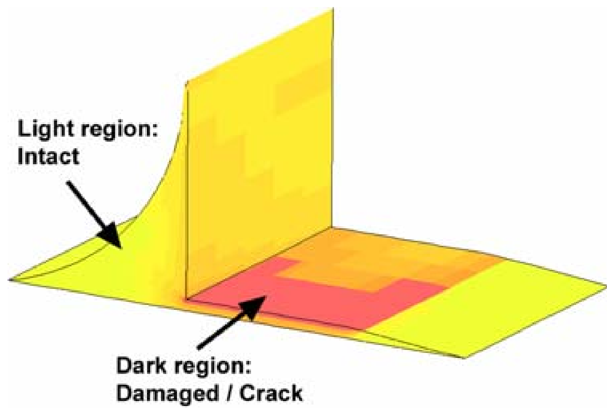

Fig. 6 Diagram showing predicted damage of surface mount resistor fillet joint after thermal cycling.

This law has been used to model SnPb[10] but the law does not capture the change in creep mechanism between low stresses and high stresses. To capture this change Wiese[18] has used a double-power model to model SnAgCu:

$$\dot{\varepsilon}_{\text{cr}} = c_1 \left(\frac{\sigma}{\sigma_{\text{n}}}\right)^{c_2} \exp\left(\frac{-Q_1}{kT}\right) + c_3 \left(\frac{\sigma}{\sigma_{\text{n}}}\right)^{c_4} \exp\left(\frac{-Q_2}{kT}\right). \tag{10}$$

The first term on the RHS of this formula represents climb controlled creep (low stress) and the second term represents glide/climb controlled creep (high stress). Constants are provided for this law to model SnAgCu.[18] An alternative and widely used method for capturing the change in creep mechanism is to use a sinh law:

$$\dot{\varepsilon}_{\text{s}} = c_1[\sinh(c_2\sigma)]^{c_3} \exp\left(\frac{-Q}{kT}\right). \tag{11}$$

Many authors have used this approach[8,17,19,20] and Darveaux *et al*. have published constants for four different alloys including SnPb and Sn3.5Ag.[4] There is spread in the published experimental creep data, differences in strain rates in the order of 10 to 100 times at a given stress and temperature are common. This can be partly attributed to different scales of sample being used by different researchers (bulk samples behave differently to joint-scale samples) as well as the fact that the strain rate is very sensitive to small changes in stress.

Primary creep/Kinematic hardening

Good lifetime predictions are made using only steady state creep laws,[4] probably because although the absolute values of strain predicted are inaccurate, the fatigue laws have been calibrated to these results and thus provide reasonable lifetime predictions. Even better predictions should be possible by modelling primary creep. Three different approaches have been found to model primary creep, the use of a *time-variable*, *isotropic hardening* and *kinematic hardening*. Of these, only *kinematic hardening* is suitable to model the solder behaviour during thermo-mechanical cycling.

Darveaux *et al*.[4] and Schubert *et al*.[19] have reported laws that make use of a *time variable* in the creep strain rate function. Darveuax's law is represented by Eq. (12), while Schubert uses a slightly more complex formula.

$$\dot{\varepsilon}_{\text{cr}} = \dot{\varepsilon}_{\text{s}}(1 + c_1 c_2 \exp(-c_2 \dot{\varepsilon}_s t)). \tag{12}$$

This may be adequate for modelling a monotonic creep test, but modelling of cyclic temperature cycling or fatigue cycling where the stress and therefore $\dot{\varepsilon}_{\text{s}}$ is constantly changing this law is not appropriate.

Cheng *et al*.[15] used the Anand model to model the primary and secondary regions of the creep curve using *isotropic hardening*:

$$\dot{\varepsilon}_{\text{cr}} = c_1 \left[\sinh\left(c_2 \frac{\sigma}{s}\right)\right]^{1/c_3} \exp\left(-\frac{Q}{RT}\right). \tag{13}$$

s is a scalar internal variable, which represents the averaged isotropic resistance to plastic flow. It changes according to the following formula:

$$\dot{s} = \left\{c_4 \left|1 - \frac{s}{s*}\right|^{c_5} .\text{sign}\left(1 - \frac{s}{s*}\right)\right\} .\dot{\varepsilon}_{\text{cr}}. \tag{14}$$

Where s^* is the saturation value of s given by:

$$s* = c_6 \left[\frac{\dot{\varepsilon}_{\text{cr}}}{c_1} \exp\left(\frac{Q}{kT}\right)\right]^{c_7} \tag{15}$$

Where c_1 in Eq. (15) is the same as c_1 in Eq. (13). The material properties for 60Sn40Pb, 62Sn36Pb2Ag and 96.5Sn3.5Ag were fitted from the conventional Darveaux model parameters. The use of an isotropic hardening law to capture the hardening during a creep test is questionable as a large part of the hardening is due to kinematic hardening, Stolkarts *et al*.[21] have even reported that isotropic hardening does not occur in SnPb solder. So although this model may capture the behaviour of a monotonic creep test it will not capture the true behaviour in cyclic testing where the load is reversed. Despite this the model has been used in a FEA analysis of chip assemblies by many researchers, often with good agreement to experimental results.[8,11,15]

Kinematic hardening

To understand how the solder behaves under cyclic loading, mechanical fatigue tests are used. During this test a cyclic strain, either shear or tensile, is imposed on a solder joint or bulk sample. The response of the solder is dependent on the elastic, creep and fatigue properties making it a useful test to validate the accuracy of constitutive laws.

Kinematic hardening laws with material parameters for SnPb have been found[22–24] which can accurately predict the behaviour shown in isothermal fatigue tests, but no laws with material parameters have been found for SnAgCu. shows experimental and simulated hysteresis loops of isothermal fatigue tests from Neu[22] predicted using the McDowell UCP (unified creep plasticity) model[25] which includes both isotropic and kinematic hardening terms.

Instantaneous plasticity

Darveaux used the following law to capture the instantaneous plasticity, which occurs in creep tests in many solders including SnPb and Sn3.5Ag:

$$\varepsilon_{\text{in}} = c_1 \left(\frac{\sigma}{E}\right)^{c_2}. \tag{16}$$

Unfortunately, this law cannot be applied to other loading conditions such as temperature cycling where the stress would be constantly changing.

Wiese has used the multilinear elastic–plastic model of ANSYS to simulate the time-independent behaviour of Sn37Pb, Sn3.5Ag and Sn4Ag0.5Cu. This was shown to accurately predict the behaviour of the solder in isothermal fatigue testing with cycle periods of 1 s to 3600 s. Wiese shows that for fatigue tests with a frequency of 1 Hz, the instantaneous plasticity dominates but for cycles with a period of 3600 s the creep dominates and the instantaneous plasticity is insignificant. Because most testing is in the order of 3600 s cycles then this would suggest that the common practice of ignoring instantaneous plasticity is justified.

It is difficult to say exactly how much of the strain is instantaneous plasticity and how much is primary creep, and even if it is worth distinguishing between the two. An alternative is to use a model such as the McDowell unified creep–plasticity model[22,25] which appears to accurately capture the behaviour of SnPb solder using isotropic and kinematic hardening and no instantaneous plasticity. This approach is preferable as it avoids the difficult-to-measure distinction of how much strain is due to instantaneous plasticity and how much is due to creep.

Although going beyond steady-state creep modelling to include instantaneous plasticity and/or hardening results in more accurate predictions of the solder behaviour, it remains unproven as to whether this will yield significantly improved lifetime predictions.

Fatigue laws

Fatigue laws are required to predict the number of cycles to fail when provided with either the accumulated effective creep strain or the strain energy from a simulation. A review of fatigue laws for solder is provided by Lee *et al*.[26] The simplest fatigue laws are the Coffin–Manson law Eq. (17) and the strain energy based law Eq. (18) used by Akay.[27]

$$\Delta\gamma = c_1 N_f^{c_2} \tag{17}$$

$$N_f = \left(\frac{\Delta W}{c_1}\right)^{1/c_2}. \tag{18}$$

If a creep law is used which outputs more than one kind of plastic strain then the fatigue law may make use of this added data to predict the contributions to lifetime from different damage mechanisms. The following equation can then be used to calculate the overall N_f:

$$\frac{1}{N_f} = \frac{1}{N_f^1} + \frac{1}{N_f^2} + \ldots + \frac{1}{N_f^n}. \tag{19}$$

Where $N_f^1, N_f^2, \ldots N_f^n$ are the number of cycles to fail predicted separately for each of the n damage mechanisms. An example of this is provided by Syed[6] where the low-stress creep and high-stress creep contributions predicted using the Wiese double power law (10) were used to predict the fatigue life:

$$N_f = \left(0.013\varepsilon_{\text{acc}}^{\text{I}} + 0.036\varepsilon_{\text{acc}}^{\text{II}}\right)^{-1}. \tag{20}$$

This law along with strain and energy based laws of the form of (2) and (18), respectively, were calibrated with a set of experimental data on assembles with four different ball pitches and sizes, three different substrate materials and three different accelerated temperature profiles using SnAgCu solder. All of the models were found to predict the characteristic lifetimes to within 25% in most cases.

Darveaux *et al*. published a methodology[4] using fatigue laws which predict the time to crack initiation and the growth rate of the primary and secondary cracks in an SnPb solder bump joint. The crack growth rates are assumed to be constant making the prediction of overall lifetime possible. This has since been updated and simplified[8] by combining the primary and secondary crack propagation rates.

Darveaux used FEA and calculated ΔW by averaging over a layer of solder 25 μm thick adjacent to the package interface where the cracks are expected to develop. This was then used in the law for predicting time to crack initiation Eq. (21), and the law for crack propagation rate Eq. (22).

$$N_o = c_1 \Delta W^{-c_2} \tag{21}$$

$$\frac{da}{dN} = c_3 \Delta W^{c_4}. \tag{22}$$

The characteristic life and failure-free life can then be calculated from Eq. (23) and Eq. (24), respectively.

$$\eta = N_o + \frac{a}{da/dN} \tag{23}$$

$$N_{ff} = \frac{\eta}{2}. \tag{24}$$

Also provided[8] is a formula to calculate the overall characteristic life based on the number of worst-case joints present in the whole system (typically only the four or eight highest-stressed joints contribute in a BGA package as this is where the vast majority of failures occur). The method predicted lifetimes with an accuracy of $\pm 2\times$ or better.

When using any fatigue law it is important to understand that it will work best under conditions similar for that which the constants were validated. Changing conditions such as chip geometry and temperature profile, and also the volume over which averaging of ε_{acc} or ΔW is performed will affect the accuracy of the lifetime predictions. The accuracy of predictions was found to be in error by a factor of up to $7\times$ when using a modelling procedure not consistent with that used in crack growth correlation.[8] The use of relative predictions is recommended[8] when there is at least one data set of measured fatigue lifetime for the package in question. The procedure is to first calculate the lifetime for the known case, then any further calculations on different cases can be 'calibrated' using the experimental data. The accuracy of relative predictions was found to be in the range of $\pm 25\%$ or better.

The *constitutive law + fatigue law* approach is an empirical method tailored to a specific experimental data set, one of the reasons that many researchers are investigating the use of damage mechanics methods is that these will hopefully remain accurate over a wider data set without having to be re-calibrated.

DAMAGE MECHANICS METHOD

In order to overcome the geometry dependence inherent in the fatigue law approach to predicting lifetime, or to allow crack paths to be predicted, damage-based constitutive laws can be used. A number of researchers have published results from such laws for SnPb solder although the approach is still in the experimental stage. A review covering several methods is provided by Desai and Whitenack.[28] Detailed validation against a large set of experimental data has not been found for any of these models and the computational cost involved is high. Given time, with the increase in computational power and continued research including more validation against experimental lifetime data, this approach may prove useful in predicting reliability.

J-integral

Ghavifekr and Michel[29] determine the relationship between the *J*-integral of a notched solder tensile test specimen and the crack growth rate and suggest that it can be applied to the prediction of growth rate in solder joints. Gu and Nakamura[30] use the method to determine the direction (but not rate) of crack propagation in a solder bump joint.

A simplistic use of the method may be to predict the crack growth rate at the small notch and assume a constant propagation speed. Darveaux has shown that in BGA bump joints the crack propagation speed is roughly constant, making this viable, however, other kinds of joints do not show constant crack speed (e.g. resistor joints[31]). A more advanced approach would be to advance the crack through the mesh by small amounts, refining the mesh to achieve a high mesh density around the crack tip, and predict the crack propagation speed at each point in the cracks life. One drawback to the use of the *J*-integral method for predicting reliability is the necessity to manually introduce small cracks (notches). It cannot predict the time or location of crack initiation. No work has been found which validates lifetime predictions using the *J*-integral method against experimental data.

Continuum damage mechanics

In this approach, the damage builds up within the bulk of the material, allowing for the modelling of regions of cracking rather than one sharp crack as in the *J*-integral method. This captures the phenomena seen in solder of small micro-cracks gradually forming within a region and coalescing to form macro-cracks.[32] An example of this approach is shown in Fig. 6 from modelling work on surface mount resistors by the authors of this paper.

A framework for implementing a continuum damage model is the DSC (disturbed state concept) which is described by Desai and Whitenack.[28] This regards a material as a composite, containing two different parts – the intact part and the adjusted/disturbed/damaged part, each of which has its own set of material properties. In the case of modelling fatigue damage, the adjusted (or damaged) part would have no resistance to shear stresses and possibly no resistance to hydrostatic stresses. The implementation of this approach is simpler than the *J*-integral methods in that the mesh does not need to be split (although refining could aid accuracy). Another big advantage is that no prior information on the location of the crack is required.

A formula based on accumulated effective strain in the solder can be used to calculate damage:

$$D = 1 - \exp\left(-c_1 \varepsilon_{acc}{}^{c_2}\right). \quad (25)$$

However, it is stated by Basaran *et al.*[33] that although ε_{acc} is often used to calculate damage, entropy is a better damage metric.

Volume averaging is necessary to avoid mesh dependence,[34] this raises the issue of the size of the volume to use. In reality, the influence of micro-crack interaction occurs over a fairly small length scale, however, if the scale is made smaller than the mesh element size then volume averaging will have no effect.

Cohesive zone

This technique models the gradual degradation in the adhesion between surfaces, making it ideal for modelling delamination at solder/pad interfaces. The method uses flat, 2D elements, called cohesive zone elements, at an interface along which a crack is expected to develop (e.g. the solder to copper pad interface, or the region of intermetallics near the top of a BGA joint). This element has a stiffness, which governs how far it separates (in normal and/or shear directions) depending on the applied stress. As the normal and/or shear separation increases or cycles then the damage present in the element increases and this in turn reduces its stiffness. This is very similar to the continuum damage approach except using 2D rather than 3D elements.

Abdul-Baqi *et al.*[35] have described a method whereby the bulk of the solder is modelled using a linear elastic law, but with cohesive zones at the solder-pad interface, the interfaces between the phases of SnPb (Pb-rich islands within a Sn-rich matrix) and even in the grain boundaries within each phase. The cohesive zones each have an associated damage parameter, which evolves according to:

$$\dot{D} = c_1|\dot{\mu}|(1 - D + c_2)^{c_3}\left(\frac{|Y|}{1 - D} - c_4\right). \quad (26)$$

An increase in the damage leads to a decrease of stiffness as governed by:

$$Y = h(1 - D)\mu. \quad (27)$$

By using cohesive zones to model four kinds of interface (solder to pad, Sn-rich phase to Pb-rich phase, Sn-rich grain boundary, Pb-rich grain boundary) the method offers a potentially very accurate description of the solder behaviour. However, a model is useless without accurate material constants and it is not mentioned how the material properties for these different cohesive zones should be found. Finding the parameters of just one kind of cohesive zone could be done reasonably by using an inverse analysis/optimisation approach, but though the approach could be applied in this case, it is likely that the relative proportion of damage occurring in each kind of cohesive zone (solder to pad, grain boundaries, etc.) will be very difficult to predict.

An interesting consequence of the cohesive zone approach is that the compliance of the cohesive zone will affect the compliance of the bulk of solder. This is an unwanted side effect and Abdul-Baqi *et al.*[35] have in their work ensured that the contribution on the joint compliance from the cohesive zones in their initial, undamaged state is negligible. The cohesive zones remain stiff under compression, thus preventing the unrealistic occurrence of overlapping crack surfaces.

A model is proposed by Yang *et al.*[36] in which a single cohesive zone element is used to model SnPb solder in a pure shear cyclic fatigue test. Strangely, the cohesive zone in this model has a thickness, therefore strains are discussed rather than separations or displacements, making it fit the description of a continuum damage model better than the cohesive zone model they describe it as. Separate laws are used to describe the damage evolution under monotonic strain Eq. (28) and cyclic strain Eq. (29).

$$\frac{\partial D_{mon}}{\partial \gamma} = 0.166\gamma^{-0.648} \quad (28)$$

$$\frac{\partial D_{cyc}}{\partial \gamma} = 0.067\gamma^{-0.543}. \quad (29)$$

This reflects the fact that it takes considerably less accumulated strain to destroy a joint under monotonic loading, than it does under cyclic loading. The way that these are used in a simulation is to use Eq. (28) to model the first half cycle, and then use Eq. (29) for all the remaining cycles. In fact the use of Eq. (28) in the first half cycle in this way is probably insignificant, given that 1000 s of cycles are typically necessary to cause failure.

Hybrid

Towashiraporn *et al.*[37] describe a method in which a damage parameter is calculated based on the accumulated effective strain according to Eq. (25), as might be done with a continuum damage approach. But unlike a continuum damage approach, the damage does not affect the stiffness of the bulk of the material; instead, the damage is monitored across the critical interface where the crack is expected to develop. When the damage at a node in this interface reaches a critical value, the connectivity at that node is released, thus creating a crack.

In order to reduce computation time, a global–local submodelling method was used. This involves first performing a simulation on a 1/8 symmetry mesh of the whole BGA package. From this, the displacements at the top and bottom interfaces of the most critical solder joint are used as boundary conditions in a more detailed simulation of the critical joint. This was found to be inadequate due to the changing response of the solder as the crack grows. To capture this, a more sophisticated method was used where the refined submodel was coupled to the whole assembly model by multiple constraints. However, this involved considerably more computational effort.

To save on computational expense, an approximate solution procedure is used which utilizes a critical disturbance range rather than a single value. The damage accumulation rate is calculated from a simulation of two cycles. This rate is assumed to remain constant and used to work out the cycle before the most damaged node will increase beyond the upper bound of the critical disturbance range. At this cycle, all nodes with damage falling within the critical range are disconnected, and the process is repeated iteratively until 70% of the interface is cracked, at which point the joint is considered to have failed. The 70% figure is chosen because it was found experimentally that only 70% of a BGA joint crack is caused by creep–fatigue and the remaining 30% is caused from shear overload. This is also supported by simulations continuing beyond 70% cracking which predict a roughly constant rate of crack growth up to 70%, followed by an exponential increase in crack growth rate. It was found that increasing the critical disturbance range from (0.75–0.85) to (0.5–0.9) did not affect lifetime predictions adversely, but it did reduce the accuracy of crack front predictions.

This is an interesting approach but as with all the damage mechanics methods is in need of further experimental validation. Also the fact that the critical surface must be specified beforehand makes this method unsuitable for modelling new or more complicated geometries.

CONCLUSIONS

For many engineers, an *analytical* method such is attractive due to its ease of use. Furthermore, both Engelmaier[2] and Clech[9] have achieved lifetime predictions with an error of less than $\pm 2\times$ with analytical approaches. However, these methods do not model the creep behaviour of the solder and are therefore limited in their capability to model a wide range of temperature conditions and geometries. The caveat restricting its use under large temperature ranges (-50–80 °C) makes Engelmaier's method unsuitable to model many accelerated thermal profiles of interest.

A *constitutive law + fatigue model* will offer more accurate results, with fewer caveats to its range of applicability. Darveaux[8] has shown that lifetimes can be predicted with an accuracy of 25% for relative predictions, and $\pm 2\times$ for absolute predictions. Syed[6] has also shown a good correlation with experimental results. The use of the Anand isotropic hardening law[15] is popular but arguably inappropriate because the kinematic hardening may have a greater influence over solder behaviour. The inclusion of kinematic hardening laws should improve prediction of solder strains but the impact on reliability predictions is at present unproven. The disadvantage of Finite Element methods is the expertise and time required to set up the analysis. This can be partly overcome by using simpler numerical methods such as Clech's SRS software,[9] but these approaches do not capture all the geometry details possible using FEA.

All of the above methods make use of an empirical fatigue law whose constants are geometry dependent, the more advanced *damage mechanics* based methods avoid the geometry dependence to a large extent by explicitly modelling the crack propagation. They offer the potential for even better accuracy of results over an even wider range of conditions, but are in need of further development and experimental validation before they can be recommended to an industry.

Acknowledgement

The work was carried out as part of a project in the Materials Processing Metrology Programme of the UK Department of Trade and Industry. Other sponsors include the EPSRC (Engineering and Physical Sciences Research Council), Prime Faraday and the NPL (National Physical Laboratory) in the UK.

REFERENCES

1 LaDou, J. Printed circuit board industry. *Int. J. Hygiene Environ. Health* (in press).
2 Engelmaier, W. (1991) Solder attachment reliability, accelerated testing, and result evaluation. In: *Solder Joint Reliability: Theory and Applications* (Edited by J. H. Lau). Van Nostrand Reinhold, New York.
3 IPC (1996) *IPC-D-279: Design Guidelines for Reliable Surface Mount Technology Printed Board Assemblies*. IPC.
4 Darveaux, R., Banerji, K., Mawer, A. and Doddy, G. (1995) *Reliability of Plastic Ball Grid Array Assembly* in *Ball Grid Array Technology* (Edited by J. H. Lau). New York.
5 Clech, J. P. (2004) Lead-free and mixed assembly solder joint reliability trends. In: *APEX Designers Summit*. IPC, Anaheim, California.
6 Syed, A. (2004) Accumulated creep strain and energy density based thermal fatigue life prediction models for SnAgCu solder joints. In: *ECTC*. IEEE.
7 Zienkiewicz, O. C., Taylor, R. L. and Zhu, J. Z. (2005) *The Finite Element Method: Its Basis and Fundamentals*. Butterworth-Heinemann, Oxford, UK.

8 Darveaux, R. *Effect of Simulation Methodology on Solder Joint Crack Growth Correlation (White Paper)*, Amkor.
9 Clech, J. P. (1997) Solder reliability solutions: a PC-based design-for-reliability tool. *Solder. Surf. Mount Technol.* **9**, 45–54.
10 Lee, S.-W. R. and Zhang, X. (1998) Sensitivity study on material properties for fatigue life prediction of solder joints under cyclic thermal loading. *Circuit World.* **24**, 26–31.
11 Pang, J. H. L., Low, P. T. H. and Xiong, B. S. (2004) Lead-free 95.5Sn-3.8Ag-0.7Cu solder joint reliability analysis for micro-BGA assembly. In: *ITHERM*. IEEE, Las Vegas.
12 Bathe, K.-J. (1996) *Finite Element Procedures*. Prentice Hall, New Jersey.
13 Pang, H. L., Xiong, B. S., Neo, C. C., Mang, X. R. and Low, T. H. (2003) Bulk solder and solder joint properties for lead free 95.5Sn-3.8Ag-0.7Cu solder alloy. In: *ECTC 2003*. IEEE.
14 Basaran, C. and Jiang, J. (2002) Measuring intrinsic elastic modulus of Pb/Sn solder alloys. *Mech. Mater.* **34**, 349–362.
15 Cheng, Z. N., Wang, G. Z., Chen, L., Wilde, J. and Becker, K. (2000) Viscoplastic Anand model for solder alloys and its application. *Solder. Surf. Mount Technol.* **12**, 31–36.
16 Dudek, R., Walter, H., Doering, R. and Michel, B. (2004) Thermal Fatigue Modelling for SnAgCu and SnPb Solder Joints. In: *EuroSimE*. IEEE, Brussels.
17 Wiese, S. and Rzepka, S. (2004) Time-independent elastic-plastic behaviour of solder materials. *Microelectron. Reliab.* **44**, 1893–1900.
18 Wiese, S. and Wolter, K. J. (2004) Microstructure and creep behaviour of eutectic SnAg and SnAgCu solders. *Microelectron. Reliab.* **44**, 1923–1931.
19 Schubert, A., Dudek, R., Döring, R., Walter, H., Auerswald, E., Gollhardt, A., Schuch, B., Sitzmann, H. and Michel, B. (2002) Lead-free solder interconnects: Characterisation, testing and reliability. In: *EuroSimE 2002*. IEEE, Paris.
20 Pang, J. H. L., Xiong, B. S. and Low, T. H. (2004) Creep and fatigue characterization of lead free 95.5Sn-3.8Ag-0.7Cu solder. In: *ECTC 2004*. IEEE.
21 Stolkarts, V., Keer, L. M. and Fine, M. E. (1999) Damage evolution governed by microcrack nucleation with application to the fatigue of 63Sn-37Pb solder. *J. Mech. Phys. Solids* **47**, 2451–2468.
22 Neu, R. W., Scott, D. T. and Woodmansee, M. W. (2000) Measurement and modeling of back stress at intermediate to high homologous temperatures. *Int. J. Plast.* **16**, 283–301.
23 Gomez, J. and Basaran, C. Damage mechanics constitutive model for Pb/Sn solder joints incorporating nonlinear kinematic hardening and rate dependent effects using a return mapping integration algorithm. *Mech. Mater*. (in press).
24 Yang, X. and Nassar, S. (2005) Constitutive modeling of time-dependent cyclic straining for solder alloy 63Sn-37Pb. *Mech. Mater.* **37**, 801–814.
25 McDowell, D. L. (1992) A Nonlinear kinematic hardening theory for cyclic thermoplasticity and thermoviscoplasticity. *Int. J. Plast.* **8**, 695–728.
26 Lee, W. W., Nguyen, L. T. and Selvaduray, G. S. (2000) Solder joint fatigue models: review and applicability to chip scale packages. *Microelectron. Reliab.* **40**, 231–244.
27 Akay, H., Zhang, H. and Paydar, N. (1997) Experimental correlations of an energy-based fatigue life prediction method for solder joints. In: *International Intersociety Electronic and Photonic Packaging Conference (INTERpack)*. ASME.
28 Desai, C. S. and Whitenack, R. (2001) Review of models and the disturbed state concept for thermomechanical analysis in electronic packaging. *J. Electron. Packag.* **123**, 19–33.
29 Ghavifekr, H. B. and Michel, B. (2002) Generalized fracture mechanical integral concept JG and its application in microelectronic packaging technology. *Sensors Actuators A: Physical.* **99**, 183–187.
30 Gu, Y. and Nakamura, T. (2004) Interfacial delamination and fatigue life estimation of 3D solder bumps in flip-chip packages. *Microelectron. Reliab.* **44**, 471–483.
31 Shangguan, D. (1999) Analysis of crack growth in solder joints.. *Solder. Surf. Mount Technol.* **11**, 27–32.
32 Wen, S. (2004) Damage based fatigue criterion for solders in electronic packaging. In: *ITHERM*. IEEE.
33 Basaran, C., Tang, H. and Nie, S. (2004) Experimental damage mechanics of microelectronic solder joints under fatigue loading. *Mech. Mater.* **36**, 1111–1121.
34 Desai, C. S., Basaran, C. and Zhang, W. (1997) Numerical algorithms and mesh dependence in the disturbed state concept. *Int. J. Numer. Methods Engng.* **40**, 3059–3083.
35 Abdul-Baqi, A., Schreurs, P. J. G. and Geers, M. G. D. (2005) Fatigue damage modeling in solder interconnects using a cohesive zone approach. *Int. J. Solids Struct.* **42**, 927–942.
36 Yang, Q. D., Shim, D. J. and Spearing, S. M. (2004) A cohesive zone model for low cycle fatigue life prediction of solder joints.. *Microelectron. Engng.* **75**, 85–95.
37 Towashiraporn, P., Subbarayan, G. and Desai, C. S. (2005) A hybrid model for computationally efficient fatigue fracture simulations at microelectronic assembly interfaces. *Int. J. Solids Struct.* **42**, 4468–4483.