CYCLIC FATIGUE OF INTRINSICALLY BRITTLE CERAMICS IN CONTACT WITH SPHERES

D. K. KIM‡, Y.-G. JUNG‡, I. M. PETERSON and B. R. LAWN§

Materials Science and Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, U.S.A.

(Received 24 March 1999; accepted 28 July 1999)

Abstract—Contact damage modes in cyclic loading with spheres are investigated in three nominally brittle ceramics, soda-lime glass, porcelain and fine-grain silicon nitride, in moist environments. Initial damage at small numbers of cycles and low loads consists of tensile-driven macroscopic cone cracks ("brittle" mode). Secondary damage at large numbers of cycles and high loads consists of shear-driven distributed microdamage ("quasi-plastic" mode), with attendant radial cracks and a new form of deeply penetrating subsidiary cone cracks. Strength tests on indented specimens are used to quantify the degree of damage. Both damage modes degrade the strength: the first, immediately after cone crack initiation, relatively slowly; the second, after development of radial cracks, much more rapidly. A fracture mechanics model describing the first mode, based on time-integration of slow growth of cone cracks, is presented. This model provides simple power-law relations for the remaining strength in terms of number of cycles and contact load for materials design. Extrapolations of these relations into the quasi-plastic region are shown to be non-conservative, highlighting the need for further understanding of the deleterious quasi-plastic mode in tougher ceramics. Comparison with static contact data indicates a strong mechanical (as opposed to chemical) component in the cyclic fatigue in the quasi-plastic region. Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.

Keywords: Structural ceramics; Fracture; Fatigue; Yield phenomena; Microstructure

1. INTRODUCTION

Fatigue properties are important in any long-term, large-scale structural applications where cyclic loads are experienced. Such properties have been measured in several ceramics using conventional pre-cracked tensile test specimens or bend bars [1–9]. A key element of the fatigue response is microstructure: fine, homogeneous ceramics, characterized by crack-size-invariant toughness values, undergo cyclic fatigue by time-integrated, chemically-enhanced slow crack growth [10]; coarse, heterogeneous ceramics, characterized by $R$-curves, undergo fatigue mainly by time-independent mechanical degradation of grain bridging or other crack-tip shielding mechanism [11–15]. The competitive roles of chemical and mechanical processes in the fatigue responses of ceramics is an issue of continuing debate.

Fatigue effects in ceramics can be even more strongly evident in Hertzian-type contact loading with spheres, where the stress field is largely compressive and of uncommonly high intensity (i.e. in excess of GPa) [16–21]. Earlier studies on damage accumulation and fracture around notches in compressive loading constitute a precedent for such effects [22–24]. Contact fatigue is relevant to bearings and engine components [21], dental restorations [25, 26] and analogous biomechanical replacements (hip joints, heart valves), and other applications where loads are concentrated. Two distinct contact damage modes have been identified [18]: (i) in homogeneous ceramics, well-defined cone cracks ("brittle" mode); (ii) in heterogeneous ceramics, microdamage within diffuse but well-defined yield zones ("quasi-plastic" mode). Again, in the brittle materials the fatigue is attributable to chemically-driven, time-independent slow crack growth, whereas in the quasi-plastic materials it appears to be governed primarily by a mechanical component, dependent on number of cycles rather than on time [27]. Both damage modes are deleterious to the remaining strength of the material. However, the mechanics of the associated fatigue processes remain obscure.

In this paper we present results of a fatigue study with spherical indenters on nominally brittle ceramics, i.e. relatively homogeneous ceramics with no significant $R$-curve, using an indentation–strength
test procedure. This procedure involves two steps: (i) cyclic contacts on specimen surfaces using a hard sphere of given radius, at prescribed frequency and maximum load; (ii) inert strength measurement of the contact-damaged specimens. The first step allows characterization of the actual damage process—the second step enables quantification of the damage severity. Glass is an ideal model brittle material because of its well-documented indentation-strength properties in single-cycle sphere loading [28–31] (see Section 2.1), as well as its availability of essential crack velocity parameters for data analysis (from Vickers indentation fracture studies [32–34]). Moreover, the transparency of glass allows for direct observations of the fracture damage from the contact tests and the response of this damage in subsequent strength tests. Other brittle ceramics selected for study include: a translucent dental porcelain, because of its brittleness and its amenability to limited subsurface observation [25, 26]; and a relatively strong bearing-grade silicon nitride [35], included to expand the range of toughness values. The bulk of the contact fatigue tests are conducted in moist environments, to highlight any chemical effect.

Using an indentation fracture analysis, we confirm that the initial stages of strength degradation are indeed consistent with chemically-enhanced slow growth of classical cone cracks, integrated over the duration of the fatigue test ("brittle" region). The analysis provides simple power-law relations for evaluating cyclic lifetimes in this region. However, beyond a critical number of cycles, a more deleterious mechanism of degradation comes into play, signaling the imminent demise of the material. This second mode is shown to be associated with the onset of distributed subsurface microdamage, with attendant development of more dangerous radial cracks ("quasi-plastic" region). Comparison of the cyclic fatigue data with static fatigue data confirms a strong mechanical component in the cyclic damage buildup in the latter region. These results highlight both the strengths and limitations of traditional fracture mechanics concepts in fatigue analysis, especially in the context of lifetime design, and foreshadow a totally dominant quasi-plastic mode of failure in tougher, more heterogeneous ceramics.

2. CONTACT FATIGUE AND STRENGTH DEGRADATION

2.1. Background

The first indentation–strength studies with spherical indenters were carried out on soda-lime silicate glass, the quintessential brittle material, in single-cycle indentation [28, 29, 36, 37]. Generally, the strength of surface-damaged glass specimens from single-cycle contacts falls off abruptly beyond a critical load for cone crack formation, and thereafter declines continuously and slowly with increasing contact load, in accordance with traditional indentation equilibrium fracture mechanics. One study with spheres in sustained static loading over prescribed contact durations ("static fatigue") on glass in water [31] reported a rate-dependent reduction of the critical load for initiation of cone cracks and subsequent propagation of the cone cracks, with corresponding steady depression in the ensuing strength levels. Those rate effects were found to be entirely consistent with time-integrated chemically-enhanced slow crack growth, in accordance with expectation for brittle solids. Another study with impacting spheres [30] reported strength degradation characteristics similar to those in static loading. In that case the main features of the degradation could be accounted for by introducing a simple relation between impact velocity and impulsive load into the indentation fracture mechanics. However, in this latter study some irregularities in the detailed strength response were noted, particularly with smaller and denser spheres—these irregularities were correlated with the appearance of plasticity-induced radial cracks from sphere penetration into the impact surface, in a manner first observed by Phillips (cited in Ref. [38]) and later quantified by Swain and Hagan [39].

Later, strength degradation studies from single-cycle sphere contacts were extended to relatively homogeneous polycrystalline ceramics [35, 40–44]. Again, cone cracking was identified as the principal mode of damage, with similar abrupt strength decrements at critical contact loading. In more heterogeneous ceramics, quasi-plastic damage also degraded the strength, initially without the abrupt falloff [27]—however, in these materials the prospect of microcrack coalescence within the damage zone was identified as a potential source of ultimate rapid strength decline in more severe indentation events [27], e.g. in single-cycle overload, multi-cycle contacts [21], or high-velocity impact [45].

2.2. Fracture mechanics analysis

Consider the application of indentation fracture mechanics to the analysis of strength degradation from cyclic contacts with spheres, Fig. 1. Begin with the assumption that the material is ideally brittle, so that the degradation is caused exclusively by failure from cone cracks. The cone fracture is assumed to be subject to slow crack growth during the cyclic contact loading according to a power-law velocity function, as previously considered in static loading [31] but now with the time integral taken over a periodic contact history. We defer, until later, consideration of failure from any secondary mode of damage.

Thus, assume that the cone crack extends in accordance with a basic power-law crack velocity
function

\[ v = v_0(K/K_0)^N \quad (K < K_0) \]  

(1)

where \( T_0 (K_0) \) is the single-value material toughness defining the upper limit of the velocity range, and \( N \) and \( v_0 \) are velocity exponent and coefficient, respectively. We consider two stages in the kinetic crack evolution in relation to the ensuing strength degradation properties: initiation, determining the threshold load at which strength undergoes an abrupt strength loss; propagation, determining subsequent continued strength loss at post-threshold loads.

2.2.1. Cone crack initiation. The radial tensile stress \( \sigma_t \) acting at the surface coordinate \( R_0 \) of a prospective cone crack from contact with a sphere of radius \( r \) at load \( P \) (Fig. 1) for a material of Poisson’s ratio \( \nu \) is [46]

\[ \sigma_t = \frac{1}{2}(1-2\nu)P/\pi R_0^2. \]  

(2)

The corresponding stress–intensity factor has the generic form [47, 48]

\[ K(c) = \sigma_0(\pi c)^{\frac{1}{2}} F(c) \]  

(3)

where \( c \) is the crack dimension measured along the downward crack coordinate \( s \) and \( F(c) \) is a dimensionless Greens function, with \( F < 1 \) for any diminishing tensile field \( \sigma_s < \sigma_0 \) at \( s > 0 \).

Suppose now that the indenter is subject to some time-dependent contact load \( P(t) \). Assume that a ring crack initiates from a surface flaw of characteristic size \( c_f \) and begins to extend downward according to equation (1). We may solve for the time to attain instability by writing \( v = dc/dt \) in equation (1) and combining with equations (2) and (3):

\[ v_0[(1-2\nu)/2\pi^{1/2} R_0^2 T_0]^N \int_0^c [P(t)]^N \, dt \]

\[ = \int_0^c dc/[c^{1/2} F(c)]^N \]  

(4)

where the instability crack size \( c = c_f \) at time \( t = t_c \) for cone crack pop-in is determinable from the limiting equilibrium condition \( K = T_0 \) in equation (3). Since the extension to depth \( c_f \) is generally small compared to the final cone crack size, we regard this phase of the crack evolution as “initiation”.

In the special case of periodic loading to peak load \( P \) over \( n \) cycles at frequency \( f \), corresponding to an indentation time \( t = nf \), equation (4) reduces to

\[ (nP)^N_c = [H(N)f/R_0/G(N)v_0][2\pi^{1/2} R_0^{3/2} T_0/(1-2\nu)]^N \]  

(5)

where subscript \( c \) denotes critical number of cycles \( n_c \) to pop-in at fixed \( P \); or, alternatively, critical load \( P_c \) at fixed \( n \). The \( G \) and \( H \) terms are dimensionless integrals [10]

\[ G(N) = \int_0^1 [P'(f/t)/P]^N \, df/t \]  

(6a)

\[ H(N) = \int_{c_f/R_0}^{c_3/R_0} dc/c/[c^{1/2} F(c/R_0)]^N. \]  

(6b)

Equation (5) may be further reduced by normalizing relative to the value of critical load for cone initiation in single-cycle contact, i.e. \( P = P_1 \) at \( n = 1 \):

\[ P_c = P_1/n_c^{1/N} \quad (\text{fixed } n) \]  

(7a)

\[ n_c = (P_1/P_c)^N \quad (\text{fixed } P). \]  

(7b)

The quantity \( P_1 \) has the functional form \( P_1(f, N, v_0, R_0, T_0, \nu) \). We will obtain an explicit relation for this function in Section 5. Strictly, \( P_1 \) also depends on \( c_f \) and \( c_p \), but these dependencies are small because of the stabilizing effect of the diminishing stress field in cone initiation, such that
the crack spends most of its lifetime in the intermediate stages of growth [38, 47].

2.2.2. Cone crack propagation. A simple analysis of fatigue for specimens containing well-developed cone cracks has been given in an earlier study [20]. We reproduce that analysis here, with minor modifications.

The stress–intensity factor for the cone crack may be approximated by the traditional Roesler relation [28, 49]

$$K = \chi P/c^{3/2}$$  \hspace{1cm} (8)

where $\chi$ is a dimensionless geometry coefficient [50]. This relation is a far-field approximation, valid in the limit $c \gg R_0$—note that it is independent of $R_0$ and sphere radius $r$. For a time-dependent load $P(t)$ we insert equation (8) into equation (1) to obtain the cone crack size $c$ after time $t$

$$\int_{c_0}^c c^{3/2} dc = v_0(\chi/T_0)^N \int_{t_0}^t [P(t)]^N dt$$ \hspace{1cm} (9)

with $c_0$ the equilibrium crack size at $K = T_0$ in equation (8) at critical time $t_c = n_c/f$ to pop-in.

For periodic loading over $n$ cycles at frequency $f$, we integrate to obtain

$$c^{3N/2+1} = [(3N/2 + 1)G(N)v_0nf^{N/2}(\chi P/T_0)^N$$ \hspace{1cm} (10)

in the limit $n \gg n_c$ and $N \gg 1$, with $G(N)$ defined in equation (6a).

The “inert” strength of a specimen containing a well-formed cone crack is given by the familiar Griffith relation

$$\sigma = T_0/\psi c^{1/2}$$ \hspace{1cm} (11)

where $\psi$ is a crack geometry parameter. Below the threshold for cone crack initiation (equations (7a) and (7b)), the strength is considered to be determined by the surface flaw size, $c = c_1$. Above the threshold, the functional dependence $\sigma(n, P)$ is determined by inserting equation (10) into equation (11). Normalizing relative to the value of the strength $\sigma = \sigma_1$ for contacts that produce cone cracks at critical load $P = P_1$ at $n = 1$, we obtain, in the limit $N \gg 2/3$

$$\sigma = \sigma_1 (P_1/n^{1/N} P)^{1/3} \quad \text{(fixed } n)$$  \hspace{1cm} (12a)

$$n = [(P_1/P\sigma_1/\sigma)^{N}] \quad \text{(fixed } \sigma)$$  \hspace{1cm} (12b)

Note that even though $n$ may vary over many decades, $\sigma$ is not expected to be sensitive to $n$ for typically large values of $N$.

As in the preceding subsection, the quantity $\sigma_1$ can be expressed in functional form, $\sigma_1(f, N, v_0, R_0, T_0, \psi)$. Again, an explicit relation for this quantity will be derived in Section 5. Strictly, $\sigma_1$ also depends on $c_0$, but this dependency is small because of the stability of the well-developed cone crack [equation (8)]—it is now the final stage of growth that is rate limiting.

Equations (10) and (11) indicate that relative to cyclic contact tests, static contact tests run at the same maximum load $P$ yield the same strengths $\sigma$ at an “equivalent” contact time

$$t_S = Gc nf$$ \hspace{1cm} (13)

noting that $G_c \leq 1$ [$0 \leq P(t) \leq P$] and $G = 1$ ($P' = P = \text{constant}$) in equation (6a) (Appendix).

3. CONTACT FATIGUE EXPERIMENTS

3.1. Materials selection

Soda-lime glass laths 90 mm × 19.0 mm × 5.65 mm were cut with a diamond scribe from as-received commercial window plate. The laths were beveled at their top edges (the edges opposite the scribed surfaces) to minimize edge failures in the subsequent strength tests, and annealed to remove any spurious residual stresses [28].

Bars 3 mm × 4 mm × 25 mm of a feldspathic porcelain in use for dental restorations (Vita Mark II*, Vita Zahnfabrik, Bad Sackingen, Germany) were ground from large blocks. Top surfaces of the bars were finished with a 1 μm diamond paste polish. The microstructure consists of a glass matrix with reinforcing sanidine crystals and residual frit particles of dimension 1–7 μm [25, 26].

Bars of the same dimension and surface finish as the porcelain were prepared from blocks of a bearing-grade silicon nitride from an earlier study (there referred to as M-Si₃N₄) [35]. In this material the microstructure consists of a bimodal distribution of elongated $\beta$ grains of length 4.0 μm and diameter 0.5 μm (≈70 vol.%) and equiaxed $\alpha$ grains of mean size ≈0.5 μm (<20 vol.%), with oxynitride glassy bonding phase (≈10 vol.%).

3.2. Contact testing

The center regions of the top surfaces were subjected to contact testing. For the glass specimens, a light pre-abrasion was first made at the prospective contact sites with 400-mesh SiC grit to provide an adequate density of flaws for ring-crack nucleation [28, 30, 31]. In the other materials, the intrinsic flaw population associated with the microstructure obviated the need for any such pre-abrasion treatment [25, 35].

The contacts were made using tungsten carbide (WC) spheres of radius $r = 3.18$ or 1.98 mm on a servo-hydraulic universal testing machine (Model 8502, Instron Corp., Canton, MA). Cyclic tests were run in repeat loading in haversinusoidal waveform up to $n = 10^7$ at frequency $f = 10$ Hz, between specified maximum and minimum loads (the latter, typically 20 N, primarily to prevent the contact from “wandering”). Comparative static tests were also made over hold times $t = n/f$ (Fig. 1), to iso-
late any mechanical contribution to the damage accumulation. In most cases the tests were conducted in distilled water to enhance the fatigue effect; other tests were conducted in laboratory air (relative humidity ≈50%).

In some cases the WC spheres themselves showed signs of deformation, especially in extended tests on silicon nitride [35], and so were rotated periodically.

3.3. Strength tests

Strength tests were conducted on the contact-damaged bar specimens in four-point flexure, with the contact surface on the tensile side. The outer and inner span dimensions used in the flexure tests were 61 and 27 mm for the glass laths, 20 and 10 mm for the other ceramics. To obtain “inert strengths” the contact sites were first dried in warm air and covered with a drop of dry silicone oil, and the specimens then broken in fast fracture (<20 ms) [33]. All specimens were examined in a low-power microscope to determine the source of failure, i.e. damage site or “natural” flaw. Any edge failures (more common in the glass specimens) were discarded from the data pool.

3.4. Damage morphology

Selected surface contact damage sites in all materials were subjected to close examination by high-power optical microscopy, both before and after flexure testing. Most observations were made in reflection mode, in Nomarski illumination, after gold coating the indented surfaces. In the transparent glass the damage sites were also examined directly in transmission mode to observe subsurface damage; and similarly in the translucent porcelain, after grinding and polishing the specimens from the back surface to a final thickness <500 μm.

Subsurface side views were also obtained by grinding and polishing cross sections down to the indentation mid-planes [51] and viewing in optical microscopy. Again, in the glass the damage was amenable to viewing in both reflection and transmission microscopy; in the porcelain, by grinding and polishing the specimens from both sides down to the mid-plane, to net thickness <500 μm. (Bonded-interface specimens [52, 53], in which specimens are pre-sectioned and restored into contact prior to indentation, were not employed here because of concern that environmental species might penetrate the interface and produce artifacts). Some of the cross sections were etched in a solution of 12% HF acid to highlight damage features.

In situ observations of the damage zone crack growth were made in selected glass specimens during the indentation–strength tests. These observations were made using a stereo-zoom microscope with backlighting, either directly through the side surface or from below using a mirror to redirect the light source (in the latter case with the upper contact surface precoated with gold to enhance reflectation). The indentation damage evolution was monitored on a video recorder.

4. RESULTS

4.1. Strength degradation data and analysis

Results of the strength degradation tests are plotted for glass/water, porcelain/water and silicon nitride/air in Figs 2–4. Data points with error bars are experimental means and standard deviations of a minimum five specimens at each specified number of cycles \( n \) or maximum contact load \( P \). Unfilled symbols represent failures from natural flaws, filled symbols failures from indentation sites; in the latter

Fig. 2. Inert strength \( \sigma \) as a function of number of contact cycles \( n \), indentation with WC spheres of radius \( r \), at maximum loads \( P \) indicated: (a) soda-lime glass in water; (b) porcelain in water; (c) silicon nitride in air. Data points are means and standard deviations, minimum five specimens per point. Filled symbols indicate failures from indentation sites—grey symbols from cone cracks, black symbols from radial cracks. Unfilled symbols indicate failures from natural flaws—box at left axis and horizontal dashed line are “laboratory” strengths (unindented specimens). Solid curves are theoretical fits to data.
case, grey symbols represent failures from cone cracks and black symbols from radial cracks (Section 4.2). Open boxes at left axis and associated horizontal dashed lines are strengths of unindented specimens. Solid curves through the data are theoretical fits.

4.1.1. Cyclic contact fatigue tests. Figure 2 shows plots of \( \sigma(n) \) data for each material/environment system, for indentations at specified sphere radius \( r \), at two selected loads \( P \). In each case the two loads are selected so as to straddle the single-cycle cone crack threshold value, \( P_1 \) (see Table 1, later). Generally, at the higher of the two selected \( P \) values the strength is substantially degraded within the first cycle—in this case, failures originate from indentation sites at all \( n \). At the lower of the \( P \) values the strength shows no degradation initially, corresponding to failures away from the contact sites at natural flaws (indicating that any initial flaw extension prior to ring-crack instability is too minor to degrade the strength); but beyond a critical number of cycles \( (n_c) \) the strength again declines above a critical number of cycles \( (n > n_c) \). Special note may be made of the data for porcelain/water at the lower load \( P = 200 \text{ N} \) in Fig. 2(b)—in this case \( P \) lies so close to \( P_1 \) that the failures are stochastically distributed between cone crack and natural-flaw failure origins.

At higher \( n \) values the strength data show a noticeable accelerating decline. This second region becomes clearly evident after only a few hundred cycles at \( P = 500 \text{ N} \) in the glass/water and porcelain/water systems; in the silicon nitride/air system it becomes apparent only after \( n \approx 10^3 \) cycles at the comparatively high load \( P = 2200 \text{ N} \). In extreme cases the glass and porcelain specimens fail during the contact test itself (arrows).

Figure 3 shows analogous \( \sigma(P) \) results for the same material/environment systems, and at the same specified sphere radius \( r \), comparing data after prolonged cycling at specified numbers of cycles \( (n > 1) \) with those in single-cycle loading \( (n = 1) \). These results confirm a significant degradation in strength from the cycling. In the glass/water and porcelain/water systems the cyclic data fall within the domain of the second strength decrement over the entire range of \( P \) (cf. Fig. 2), and so tend to fall beneath the predicted curve; conversely, in the silicon nitride/air system the cyclic data remain within the bounds of the first strength decrement at all \( P \).

The solid curves in Figs 2 and 3 are best fits of equations (7a), (7b), (12a) and (12b) to the data in the initial degradation region corresponding to failure from cone cracks (i.e. grey symbols), using \( N \), \( P_1 \) and \( \sigma_1 \) as parameters. Vertical solid lines indicate critical \( n_c \) or \( P_c \) values for cone crack initiation [equations (7a) and (7b)]; the solid curves at \( n_c > 1 \) and \( P_c > 1 \) represent strength degradation functions for failure from these cone cracks [equations (12a) and (12b)]. For soda-lime glass, \( N = 17.9 \) is prede-

determined from previous (Vickers indentation) studies [32–34], with just \( P_1 \) and \( \sigma_1 \) as adjustable parameters; for the other materials, \( N \), \( P_1 \) and \( \sigma_1 \) are all treated as adjustable parameters. The fitted curves are consistent with the main data trends in the cone crack region, within the data scatter.

4.1.2. Cyclic vs static fatigue tests. Figure 4 compares \( \sigma(t) \) data for cyclic and static tests, again for the same material/environment systems, with the time scale for the cyclic data determined as \( t = G_0 n f \) according to equation (13). Chosen values of \( n \) and \( P \) are such that degradation occurs within the first cycle. In the region of the first strength decrement the data overlap, within the scatter bars. However, in the region of the second strength decrement the cyclic data diverge markedly below
the static data, indicating the onset of a significant mechanical degradation component in the fatigue.

Equation (12a) is used to generate the solid curves $s_C(t)$ and $s_S(t)$ in Fig. 4, using the calibrated parameters in Table 1 along with appropriate evaluations of $G_C(N)$ (Appendix).† Using equation (13) to plot the time axis for the static data reduces the $s_C(t)$ and $s_S(t)$ data to universal curves in the region just beyond the first strength decrement.

Note that whereas the $s_S(t)$ function extrapolates through the data over the complete time range, confirming a predominantly chemical fatigue component in the static fatigue, the $s_C(t)$ function again declines to lower values beyond the second decrement.

4.2. Damage morphology

Figure 5 shows surface and section optical micrographs of indentation damage sites in the three test materials, comparing damage after single-cycle ($n = 1$) and multi-cycle ($n \gg 1$) loading. The glass [Fig. 5(a)] and porcelain [Fig. 5(b)] are viewed in transmitted light, the latter after back-thinning; the silicon nitride is viewed in reflected light [Fig. 5(c)]. The loads and sphere radii chosen here highlight the transformation in damage mode (cf. Figs 2 and 3): from simple, well-defined cone cracks just outside the contact in the single-cycle state to accumulated damage below the contact, with attendant radial cracks as well as incipient material removal, in the multi-cycle state.

Figure 6 shows reflection optical micrographs of failure origins at indentation damage sites in (a) glass and (b) porcelain strength specimens, in sequences of increasing $n$. The loads and sphere radii are again chosen to highlight the transformation in damage mode. At $n = 1$ the fracture paths occur well outside the contact circle, consistent with failure from cone cracks. At intermediate $n$ the fracture paths tend to move closer to the contact circle (more evident here in the porcelain), as damage builds up within the contact area and begins to compete with the cone crack for dominance as a starting flaw in the flexural field. In the case of glass the fracture origin occurred outside and inside the contact circle with approximately equal probability at $P = 500$ N and $n \equiv 10^3$—cf. Fig. 3(a). At large $n$ the fracture paths intersect the indentation, indicative of failures from radial cracks. Similar observations have been reported previously for the silicon nitride [44] and are not reproduced here: suffice it to say that in this material the failure initiates from cone cracks over the data range except at the highest value of $n$ and $P$ in Fig. 2.

In situ observations of the indentation damage sites during the strength tests provided more definitive confirmation of the failure origins in the transparent glass specimens. In specimens with light damage the failure initiated at the base of the cone cracks and spread laterally and upward, forming the characteristic inward-pointing cusp-like trace outside the contact on the upper surface seen in Fig. 6 at $n = 1$. This kind of failure occurred rapidly, with little extension before instability. In specimens with heavy damage, the failure was confirmed to initiate from the edge of a favorably oriented radial crack (relative to the tension axis). The radial cracks extended outward $\approx 20\%$ in dimension prior to failure, indicating a stabilizing

†Actually, including the factor $G_C(N)$ in equation (13) causes only minuscule shifts of the cyclic fatigue curves along the logarithmic time axis in Fig. 4, so that omitting it would still result in almost indistinguishable cyclic and static functions.

Fig. 4. Comparison of cyclic and static contact fatigue, indentation with WC spheres of radius $r$, at maximum loads $P$ indicated. Inert strength as a function of number of contact cycles for: (a) soda-lime glass/water; (b) porcelain/water; (c) silicon nitride/air. Equivalent time axis for cyclic test data computed as $G_C N f$. Data points are means and standard deviations, minimum five specimens per point. Circles are cyclic data, triangles static data—grey symbols from cone cracks, black symbols from radial cracks. Box at left axis and horizontal dashed line are "laboratory" strengths (unindented specimens). Solid curves are theoretical predictions from cone crack model.
factor in the crack driving force. Examination in crossed polars revealed the presence of considerable residual stress around the more heavily damaged sites. Such a stabilizing influence from residual contact stresses has been well documented in earlier Vickers indentation studies [54–56].

Analogous surface observations of damage sites at high loads $P$ in fast single-cycle contact tests or

<table>
<thead>
<tr>
<th>Material</th>
<th>$N$</th>
<th>$P_1$ (N)</th>
<th>$\sigma_1$ (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soda-lime glass</td>
<td>17.9</td>
<td>228</td>
<td>90</td>
</tr>
<tr>
<td>Porcelain</td>
<td>30</td>
<td>205</td>
<td>120</td>
</tr>
<tr>
<td>Silicon nitride</td>
<td>22</td>
<td>1900</td>
<td>650</td>
</tr>
</tbody>
</table>

**Table 1. Fracture parameters for each material**

Fig. 5. Optical micrographs of Hertzian indentation sites, comparing damage after single-cycle ($n = 1$) and multi-cycle ($n \gg 1$) loading: (a) soda-lime glass/water ($P = 500$ N, $r = 3.18$ mm); (b) porcelain/water ($P = 500$ N, $r = 3.18$ mm); (c) silicon nitride/air ($P = 2200$ N, $r = 1.98$ mm). Surface views: glass and porcelain viewed in transmitted light (porcelain after back-surface thinning), silicon nitride in reflected light (Nomarski contrast, after gold coating). Note appearance of radial cracks at contact periphery after cycling.
at equivalent test durations \( t \) in extended constant load tests revealed no obvious signs of radial cracking in any of the materials. However, notwithstanding the fact that the predicted strength function in equations (12a) and (12b) fits the data over the entire range of \( P \) at \( n = 1 \) in Fig. 3 and the static data over the entire range of test time \( t \) in Fig. 4, the failure paths in the glass and porcelain did tend to intersect the contact zone at large \( P \) and \( t \) (black symbols in Figs 3 and 4), suggesting the existence of some rate effect in the radial crack initiation process.

Half-surface and section optical micrographs of damage zones in the glass and porcelain in the heavy damage region are shown in Fig. 7, for tests at load \( P = 140 \) N. The sections are viewed in transmitted light, the porcelain after thinning from both sides down to the mid-plane. Secondary cone cracks propagate steeply downward from well within the contact circle. The surface traces of these...
secondary cracks appear to correlate with the inner radius of the Hertzian contact, associated with the minimum load used in the cyclic tests—inserting $P_{\text{min}}/P_{\text{max}} \approx 0.14$ (20 N/140 N) into the classical Hertzian relation between contact radius $a$ and load $P$, $a \propto P^{1/3}$ [58] yields $a_{\text{inner}}/a_{\text{outer}} \approx 0.5$, consistent with the relative scale of the annular zone for the indent on glass in Fig. 7(a). Interestingly, these deep cracks do not appear to participate in the subsequent flexural failure, possibly because they extend down into the compressive zone of the flexure field in the strength testing; however, they may be the cause of failure during extended contact fatigue testing at high $P$ and $n$ (arrows, Fig. 2). No such analogous subsidiary cone cracks were observed in comparative static loading.

5. DISCUSSION

5.1. General

We have shown that intrinsically brittle materials, specifically soda-lime glass, porcelain and silicon nitride, are susceptible to fatigue in contacts with spheres. The strength degradation plots in Section 4.1 usefully quantify this fatigue, thereby providing a basis for materials evaluation in applications involving concentrated oscillating loads. For this purpose, it is convenient to condense data from such plots onto a “master diagram”—accordingly, in Fig. 8 we replot some of the curves from Fig. 2, along with companion curves for a glass-infiltrated alumina and yttria-tetragonal-polycrystal zirconia from a study on dental ceramics [26]. It is clear that materials like silicon nitride and zirconia, by virtue of their intrinsic toughness, occupy a high position in such a diagram, and present themselves as superior candidates for bearing materials. Nevertheless, even the strongest and toughest materials are susceptible, to a greater or lesser extent, to some degree of degradation. Moreover, there are other considerations, such as cost and aesthetics (e.g. dental restorations), that may in some cases constrain selection to the lower portions of the plot. We shall return to the issue of materials design later.

We have also demonstrated the existence of at least two modes of damage in nominally brittle ceramics. The first mode is classical cone fracture, driven in moist environments by chemically-enhanced slow crack growth. Cone fracture accounts for the first stage of degradation in the strength plots—the
“brittle” region. This region covers the bulk of the data range for the silicon nitride in Figs 2–4, but only a portion of the range for the glass and porcelain. The second mode involves development of a subsurface damage zone, predominantly driven by a mechanical degradation process (but also susceptible to chemical enhancement), with attendant radial cracks. This mode accounts for the more deleterious second stage of strength degradation—the “quasi-plastic” region. We consider each of these damage regions in more detail below.

5.2. Brittle region

Conventional fracture mechanics combining a power-law velocity function with stress-intensity factor functions for cone cracks adequately describes the brittle region of strength degradation. Within the limits of this region the fracture mechanics analysis offers relatively simple power-law relations for engineering lifetime design in terms of the important contact fatigue variables, indentation load \( P \) and number of cycles \( n \). Thus, equation (7b) defines the critical number of cycles \( n_c \) to initiate cone cracking at fixed maximum cyclic contact load \( P < P_1 \), load to initiate crack at \( n = 1 \):

\[
n_c = (P_1 / P)^N. \quad (14)
\]

Similarly, equation (12b) defines the number of cycles \( n_\sigma \) at supercritical load \( P > P_1 \) to cause failure under the action of a subsequent applied stress \( \sigma < \sigma_1 \), strength at \( n = 1, P = P_1 \):

\[
n_\sigma = [(P_1 / P)(\sigma_1 / \sigma)]^N. \quad (15)
\]

Where an application demands the avoidance of any kind of macroscopic crack initiation over a prescribed lifetime (e.g. ceramic bearings), equation (14) is appropriate. On the other hand, where limited cracking might be tolerated, and the requirement instead is sustenance of some minimum stress level over the same prescribed lifetime (perhaps some dental restorations), equation (15) applies. Clearly, the lifetimes \( n_c \) and \( n_\sigma \) are sensitive to values of operating contact load \( P \) and subsequent stress \( \sigma \), especially for materials with large \( N \).

From equations (14) and (15) it is apparent that the material requirements for long lifetime in cyclic contact are for large \( N, P_1 \) and \( \sigma_1 \). This raises the issue of the dependence of the latter two quantities on intrinsic material parameters, principally on toughness \( T_0 \). We may determine this dependence by eliminating the scaling dimension \( R_0 \) from the formulations in Section 2.2. Thus, write \( R_0 = \beta a \) where \( \beta = 1.0–1.25 \) is a dimensionless constant and \( a \) is contact radius (geometrical similarity) [35, 57]. From Hertz theory for elastic contacts [46, 58], \( a^3 = 4kPr^3/E \), with \( r \) the sphere radius, \( k \) a dimensionless constant of order unity (dependent on relative elastic properties of indenter and specimen), and \( E \) Young’s modulus. Then equation (5) at \( P = P_1, n = 1 \) yields

\[
P_1 = Ar \hat{T}_0^3/E
\]

with \( A = [16\pi k \beta^2 / 3(1 - 2\nu^2)](H/R_0/G_0)^{2/3} \), a dimensionless quantity, approximately constant (note that the last term in parentheses retains \( R_0 \), but is very slowly varying for large \( N \)). [Equation (16) is a statement of the celebrated Auerbach’s law, \( P_1 \propto r \) [38, 47, 59, 60].] Similarly, equations (10) and (11) at \( \sigma = \sigma_1, P = P_1, n = 1 \) yield

\[
\sigma_1 = (BT_0^3 E/r)^{2/3} \quad (17)
\]

with \( B = [1/(4\pi \beta)](2r^3/3N)^{1/3} \), another dimensionless term, again approximately constant. Substitution of equations (16) and (17) into equations (14) and (15) yields \( n_c \propto (P_1 / \sigma_1)^N \propto (T_0 / E)^N \), \( n_\sigma \propto (P_1 / \sigma_1)^N \propto T_0^N \). Thus the principal requirement in this region is to maximize toughness. A plot of \( \sigma_1 P_1^{1/3} \) vs \( T_0^{1/3} \) highlights the dependence of \( n_\sigma \) on \( T_0 \). Note the appearance of \( r \) and \( E \) in \( n_\sigma \) but not in \( n_c \), reflecting a diminishing sensitivity to local indentation conditions in the transition from near-field to far-field solutions [28, 44].

Quite apart from its restriction to the brittle region, the current fracture mechanics analysis is subject to certain limitations. As indicated in Section 2.2, there are several assumptions and approximations in the analysis, so the dependencies in equations (14) and (15) are not exact; on the other hand, the slow variations of \( \sigma_1(n) \) and \( \sigma(P) \) in Figs 2–4, together with the data scatter, might tend to obscure any such deviations. Statistical consideration of the data scatter itself is also an important element not considered here—such considerations become especially important in any extrapolations to long lifetimes [61]. Also, we have considered the effect of cyclic loading only in the contact process, as a function of time, and not in the surface contact process, as a function of indentation depth.
and not in the subsequent strength test—here, excluding rate effects in the latter test enables the contact fatigue process to be studied in isolation. Admission of slow crack growth in the strength test would simply be to depress the curves in Figs 2–4.

But the most severe limitation on the fracture mechanics model is the incidence of the second, quasi-plastic mode with its associated radial cracks.

5.3. Quasi-plastic region

The cone crack model is not equipped to account for the second strength decrement associated with the incidence of the quasi-plastic mode of damage at large $n$ or $P$. This second region is much more deleterious to strength than cone cracking—its onset signals the imminent end of useful lifetime (as well as of the beginning of gross material removal). Consequently, data extrapolations using simple cone fracture mechanics analysis may result in grossly non-conservative strength predictions. Ignoring the quasi-plastic mode in the materials design for any application involving repeat contact loading carries a large element of danger.

The onset of quasi-plasticity occurs subsurface, where the yield stresses have their maximum values [38, 39, 52, 53, 62–64]. Quasi-plasticity occurs earlier relative to cone cracking with diminishing indenter sphere radius, as cone cracking becomes suppressed and the indenter penetrates more easily; i.e. a “blunt–sharp” indenter transition [25, 38, 47, 50, 65]. In the limit of ideally sharp indenters, e.g. Vickers diamond pyramids, radial cracks become the dominant mode of indentation fracture and lead to accelerated strength degradation [28, 29, 66]. In relatively brittle ceramics this subsurface damage may be obscured below preceding cone cracks—it is consequently not always easy to detect in its early stages, making it even more insidious.

Ceramic microstructure has been repeatedly alluded to as a key element in quasi-plastic damage. We have focussed our attention here on the more homogeneous ceramics, i.e. ceramics without significant $R$-curves. Even the most brittle of these, glass, is to some degree susceptible to quasi-plastic damage. Introducing some level of crystallinity into the glass, as in porcelain, is expected to enhance the quasi-plasticity mode at the expense of cone cracking, leading to higher susceptibility to quasi-plastic microdamage accumulation—it has been well demonstrated in glass-ceramics [17, 25, 53, 67] and silicon nitrides [35, 44, 68] that the damage mode undergoes a profound brittle–plastic transition with increasing microstructural heterogeneity. In the toughest ceramics cone fracture may be suppressed altogether, even in low-load, single-cycle contact. One consequence of this is that material toughness no longer remains a useful indicator of damage resistance—indeed, contact fatigue data on the above-mentioned glass-ceramics [67] and silicon nitrides [21] confirm that toughening procedures, by enhancing the tendency to quasi-plastic deformation, substantially accelerate degradation rates at high $n$. Some early modeling of contact-induced quasi-plastic damage, elucidating the mechanics of single microdamage “shear faults” [18] and the subsequent coalescence of these faults, has been attempted [27, 69–71]. Extension of such modeling to contact fatigue, especially in the context of microstructure, remains a key area for future research.

Mention may be made of the largely mechanical nature of the quasi-plastic mode in the fatigue, as highlighted by the accelerated decline in strength (Fig. 4) and the corresponding development of contact surface damage (Figs 5 and 6) in cyclic relative to static loading, notwithstanding the intimation that quasi-plasticity intensifies even in static loading at long enough contact times (Fig. 4). As described in Section 4.2, the surface damage in cyclic loading occurs predominantly in an annular region delineated by the outer and inner contact circle at the maximum and minimum contact loads. It would appear that the damage is associated in some manner with a periodically expanding and contracting contact circle, augmented by the access of environmental fluid into the recontact zone. The role of the fluid at the oscillating indenter/specimen interface and any associated hydraulic pressure, especially in relation to the deeply penetrating cone-like cracks shown in Fig. 7, remains to be determined. (Recall that these deep cracks do not participate directly in the subsequent flexural failure.)

The fact that quasi-plasticity may also evolve in static contact, albeit much more slowly than in cyclic, and in single-cycle contact at high overloads, suggests some complex interplay between mechanical and chemical forces in the fundamental shear fault processes [50]. These phenomena are currently under further study.

5.4. Role of flaw state

A final word on the surface or microstructural flaw state on crack initiation is in order. Whereas starting flaw size $c_f$ is usually considered to be a critical element in the mechanical integrity of brittle solids, this is less true in indentation fracture, especially cone fracture. The rapidly diminishing tensile stress field acting along the cone crack path stabilizes the fracture [equation (2)], rendering the crack initiation insensitive to $c_f$. This insensitivity lies at the root of Auerbach’s law [equation (16)] (and of the corresponding breakdown of a critical stress concept for cone fracture) [28, 38, 47, 50, 72]. It is our justification for omitting $c_f$ (and $c_f$) from the functional dependence of the quantities $P_f$ and $\sigma_f$ in equations (7a), (7b), (12a) and (12b). This is not to imply that starting flaws are unnecessary for the nucleation of the cone cracks in the first place—whereas the flaw size is not so important, we still must guarantee a sufficient density of flaws in the vicinity of the expanding contact circle.
Hence the reason for abrading the glass surfaces in our experiments (Section 3.2).

The role of flaw state in quasi-plastic damage and ensuing radial cracking and secondary (deeply penetrating) cone cracking is less well documented. Whereas penetrating indenters are capable of generating their own starting flaws from within the quasi-plastic zone [73], there is evidence from earlier sphere impact studies [30, 45] that such cracking may be strongly augmented by pre-existing surface damage. This is another area that warrants further study.

6. CONCLUSIONS

1. Contact damage modes in cyclic loading with spheres have been investigated in relatively brittle ceramics in moist environments. Two modes, conventional tensile-driven cone cracking (“brittle” mode) and shear-driven microdamage accumulation (“quasi-plastic” mode), have been described. The latter generates radial cracks and deeply penetrating secondary cone cracks.

2. (Inert) strength degradation resulting from the foregoing damage has been quantified for the test materials. Initial strength losses are caused by the cone cracks; subsequent and much more deleterious losses are caused by the radial cracks.

3. A fracture mechanics model based on time-integration of slow growth of cone cracks has been presented. This model describes the strength characteristics in the brittle region, providing simple power-law relations for strength in terms of number of cycles and contact load for materials design.

4. Extrapolations of the above fracture mechanics analysis into the quasi-plastic region are shown to be highly non-conservative. Comparison with static contact data in the latter region implies a strong mechanical (as opposed to chemical) component in the cyclic fatigue. The quasi-plastic mode needs further study, especially in the context of ceramic microstructure but also in relation to flaw state.

Acknowledgements—The authors acknowledge valuable contributions from Seung-Kun Lee, Kee Sung Lee and Jeong Gu Yeo, for assistance in collecting some of the data. This work was funded in part by the National Institute for Standards and Technology (NIST internal funds) and in part by grants from the National Institutes for Dental Research (NIDR PO1 DE10976) and the Korea Science and Engineering Foundation (KOSEF).

REFERENCES

APPENDIX A

Calculation of Function $G_c(N)$ for Cyclic Contact Loading

We seek to calculate the function $G(N)$ in equation (6a) for cyclic loading over one time period $0 \leq t \leq 1$ at frequency $f$:

$$G_c(N) = \int_0^1 \left( \frac{P'(ft)}{P} \right)^N dt$$

(A1)

with $N$ the crack velocity exponent. The cyclic contact load has a haversinusoidal waveform:

$$P'(t) = \frac{P}{2} \left[ 1 + \sin(2\pi ft - \frac{\pi}{2}) \right]$$

(A2)

Figure A1 plots the function $\left[ P'(ft)/P \right]^N$ for selected values of $N$. A higher crack velocity exponent "sharpens" this function. A plot of $G_c(N)$ is given in Fig. A2. This function has a maximum value of 0.5 at $N = 1$ and diminishes...
monotonically with increasing $N$. It may be represented to good approximation by the empirical relation $G_C = \frac{0.50}{N^{0.47}}$ over the range $1 \leq N \leq 1000$ in Fig. A2. Note that for static contact load $P(t) = P = \text{constant}$, equation (6a) yields $G_s = 1$ as our reference baseline.

Fig. A2. Plot of $G_C(N)$ from equation (A1), for haversinusoidal waveform in equation (A2) and for selected $N$. 