

Reminiscences on Four Decades of Struggle and Progress in Softening Damage and Size Effect

By Zdeněk P. Bažant¹

Abstract: Far from attempting a comprehensive and balanced account of the evolution of the field of distributed softening damage, quasibrittle fracture and its scaling, the author reminisces in an anecdotal form on his and his collaborators' efforts to contribute to the resolution of various problems.

Four decades ago, when I started working as a bridge engineer in Prague, the material models and methods of structural analysis for concrete, as well as rocks, ice, fiber composites and other materials that are now called quasibrittle, were from today's viewpoint rather simplistic. The progressive softening damage due to distributed cracking was either ignored or misrepresented as plasticity. The size effect on the strength and ductility of structures made of these materials was mostly unknown or disregarded, and when its existence was admitted it was perceived as solely statistical and thus appropriate for being buried in the safety factors.

The advent of computers and the finite element method changed everything. It put the entire field of material modeling in flux. In January 1969, I arrived at Boris Bresler's invitation to the University of California in Berkeley, to join what was at that time the leading center for structural engineering. As I learned, the finite element method captivated everybody's mind. Being already curious about the fracturing of concrete and laminates, thanks to my previous fellowship under L'Hermite, my mentor in Paris, I became fascinated by Rashid's idea to simulate by finite elements the cracking in a nuclear reactor vessel in a smeared manner—through the stress-strain relation.

I discovered, however, that all this excitement in Davis Hall was not shared in Etcheverry Hall across the street. I think I was the only one from Davis hall to frequent the seminars in that separate world of the mechanics department. Professor Naghdi, the chairman and one of the gurus of continuum mechanics at that time, noticed me and asked: "By the way, what's your interest?" "Strain-softening, to model distributed cracking of concrete and rock", I replied. Then, in a mildly sarcastic tone, he advised me: "Young man, taking such a controversial path, you will never achieve tenure. A tangential moduli tensor whose matrix is not positive definite is not a sound concept. Materials with such a property do not exist; they would be unstable and could not propagate waves." Soon I realized that Prager and other continuum mechanics giants thought likewise, and that there were classical works beginning with Hadamard to support this viewpoint. So I decided to play it safe and focused my efforts at Berkeley on thermodynamics of creep and nano-pore water in concrete, another big issue for concrete reactor vessels.

In the faculty group at Northwestern University, which I joined after Berkeley, the major thrust of research was not finite elements. Rather, it was fracture mechanics. So I tried to educate myself on the subject. In 1972, while browsing in our library, one paper in the Indian Concrete Journal caught my eye. P.F. Walsh, a young Australian, then unknown to me, reported remarkable fracture experiments on notched concrete beams. They revealed a strong size effect. But that size effect did not follow a power law and thus conflicted with the Weibull statistical theory, then reigning supreme. Moreover, the observed size effect was much too strong to be explained by that theory. So it was clear that brittle heterogeneous materials called for a different size effect theory.

Being at Northwestern, I had the benefit of an outstanding solid mechanics group, which was young, competitive and collaborative. The group spread over several departments, and there were lively mechanics seminars to bring us together on Fridays, followed by parties long into the night. In one memorable seminar, with the provocative title 'Red Herring and Sundry Fish in Continuum Mechanics' (which actually attracted some biologists), Ronald Rivlin ridiculed some anointed principles of continuum mechanics, such as the 'principle of equipresence'. For me the effect was to lose my fear of delving into the taboo of strain softening (were I still behind the Iron Curtain, it would have been different, because wasting the resources of socialism on a study of strain softening was actually

¹Walter P. Murphy Professor of Civil Engineering and Materials Science, Northwestern University, Evanston, Illinois 60208; z-bazant@northwestern.edu.

prohibited by an omnipotent committee of the Academy). Besides, I realized that physicists were not afraid of models akin to strain softening (such as the classical van der Waals phase transition in the pressure-volume diagram of steam).

Jim Rice came to Northwestern, and his enthusiastic seminars, as well as the discussions that followed, were an eye-opener. I became intrigued by his idea (developed with his student John Rudnicki who later became my colleague at Northwestern) that the cause of localization of plastic strain into shear bands is the geometrically nonlinear effect of finite strain, and by his work with Andrew Palmer showing the size effect implied by their Mode II cohesive crack model for slides in overconsolidated clay. I came to understand that inelastic strain tends to localize, and that some sort of characteristic length of the material would have to govern it.

By 1975, efforts to analyze stability and bifurcation based on the second-order work, partially inspired by some earlier bifurcation studies of softening by Giulio Maier in Milano, eventually led me to realize that smeared cracking, or strain softening, makes sense only if the width of a localizing diffuse cracking band in a continuum is limited to a certain minimum, representing a material property, termed the characteristic length. Naturally, this further implied a deterministic size effect not unlike that discovered experimentally by Walsh.

By lucky coincidence, in 1974 Stan Fistedis invited me to draft a research proposal to AEC on problems of failure analysis of concrete vessels and containments under various hypothetical scenarios of nuclear accidents in a liquid-metal-cooled breeder reactor. Once granted, that naturally led to my long-term consulting for Stan's group at Argonne National Laboratory. "How should we extrapolate from normal-scale laboratory specimens to these very large, and politically very sensitive, structures?", Stan quizzed me. It was necessary to somehow realistically take into account distributed cracking. For that, it was inevitable to postulate strain-softening, which gives rise to a deterministic size effect.

In 1976, Arne Hillerborg, in his unpretentious manner, delivered a memorable seminar at Northwestern in which he presented his fictitious crack model for concrete—an adaptation of the cohesive crack model that evolved from Barenblatt's 1979 model for fracture on the atomic scale and whose relation to fracture energy was previously clarified by Rice. Arne showed that, unlike ductile-brittle materials, non-ductile softening materials such as concrete allow using the cohesive crack model not only for the propagation of an existing crack but also for the initiation of a crack anywhere in the material. He demonstrated by finite elements that his model, devoid of any strength randomness, could capture the size effect on the flexural strength of unnotched concrete beams, in agreement with experiments. Subsequently, at a dinner in the great 'Cafe Provencal', Arne and I argued gently about the relative merits of his model and the crack band model.

In retrospect, these arguments appear to have been unnecessary, because both models turned out to be essentially equivalent and give about the same results, the choice between them being principally a matter of convenience. The cohesive crack model is convenient for exploiting elastic solutions and connecting to linear elastic fracture mechanics (LEFM) concepts, while the crack band model is convenient for implementation in existing finite element programs and can simulate wide non-localized zones of cracking.

An inevitable conclusion from the stability and bifurcation studies of strain-softening constitutive models was that finite element analyses utilizing such models, which became enormously popular during the 1970s, were unobjective in the sense of mesh bias, particularly their spurious sensitivity to mesh refinement. Upon refining the mesh size to zero, such analyses would predict the failure to occur with zero energy dissipation. This turned out to be an inflammatory proposition which elicited vociferous opposition from structural engineering practitioners. This was the situation, for example, at the SMiRT conferences in London, San Francisco, Berlin and Paris and various ASCE meetings. At one of them, in Waterloo, I met Luigi Cedolin, who agreed with me. In 1977, he came to Evanston for his sabbatical, to help me with his insight and persistence in formulating and verifying the initial energy-based crack band model. Subsequently, with the invaluable help of Byung Oh, this was extended to a damage constitutive model with progressive strain softening. Bhat, S.-S. Kim, Tsubaki, J.-C. Chern and F.-B. Lin provided further valuable assistance in extending plasticity to develop refined constitutive models for gradual strain softening of concrete. Conceptually, the phenomenon of fracture localization was at Northwestern also clarified by stability analysis of a system of parallel thermal or drying cracks, pursued jointly with Ohtsubo who came from the University of Tokyo.

The crack band model, however, was not seen as a panacea by continuum mechanicians. At Atluri's NSF workshop at Stone Mountain in 1982, Ivan Sandler objected: "If your crack band model does not permit the element size to be reduced zero, how do you define convergence? What is the mathematical connection to the boundary value problem, which would be ill-posed anyway?" At that workshop, the

preceding one at General Butler Park in 1980, and other meetings, similar comments were made by Read, Hegemier, Valanis, Leckie and others. I had to concede that the boundary value problem had to be somehow regularized.

At Northwestern, I debated possible remedies with my friend and rival Ted Belytschko, and eventually we joined forces under our shared AFOSR and DNA grants. This led in 1983 (with the assistance of T.-P. Chang) to the nonlocal model for strain-softening, along with its second-gradient approximation, the idea of nonlocality in solids being borrowed from the earlier elastic and plastic-hardening theories of Eringen, Edelen, Kröner and others. But the first version (termed ‘imbricate’) turned out to be cumbersome to program and prone to diverge in periodic zero-energy instability modes, which had to be suppressed by an artificial parallel coupling.

A better answer materialized soon through a ‘gift’ from France—Gilles Pijaudier-Cabot came to satisfy his military service as my doctoral student. I had to report on his progress to a French colonel, and the progress was good. Once in 1986, Gilles rushed into my office, exclaiming: “Since the nonlocality of total strain is what causes all the trouble, why don’t we treat as nonlocal only the inelastic strain?” And that worked (recently, though, the numerical microplane studies of di Luzio at Northwestern revealed that it is nevertheless useful to resurrect the total strain in a kind of ‘over-nonlocal’ form, provided that it is used to control only the softening yield limits, and not the elastic strains).

The size effect law for structures developing large cracks prior to the maximum load, formulated in 1983, was initially hard to defend, even at home. For instance, after my seminar, my colleague and friend Toshio Mura commented in his caustically friendly manner: “Zdeněk, your size effect formula is *too simple*. It has to be an oversimplification at best. And your derivation was not rigorous at all!” It was hard to counter that. So, in my group of able assistants, we worked strenuously to check this law by numerical simulations using finite elements as well as micromechanical random particle models, and to verify and calibrate it experimentally.

The availability of a size effect formula made it possible to determine (with the assistance of Pfeiffer and later Kazemi) the fracture energy and other fracture parameters by measuring only the peak loads, which is easier and thus less prone to errors than post-peak measurements. In the mid-1990s, my esteemed friend and collaborator Jaime Planas, together with Elices and Guinea, extended this idea to determining solely from maximum loads the initial softening slope of the softening curve of cohesive crack (or crack band). This slope, unlike the tail of that curve, is the important parameter for calculating the load capacities of structures.

My interaction with these first-rate researchers in Spain (which began in 1985 with Manuel Elices’ idea of tapping the funds obtained by Spain for renting their air bases to the Americans, funds that were required by the treaty to be channeled into joint research) bore much fruit—for example the verification of the size effect method of fracture testing and clarification of the effects of temperature, moisture content and loading rate on concrete fracture characteristics—efforts to which also my assistants J.K. Kim, Phil Pfeiffer and Ravindra Gettu contributed mightily.

On the theoretical side, more rigorous derivations of the scaling law were attempted as well, based on the asymptotics of the cohesive crack model and on the asymptotic matching between the equivalent LEFM and plasticity. The efforts succeeded. They clarified the limits of applicability, and led to ramifications for brittle shear and compression failures, borehole breakout in rock, and triggering of snow avalanches.

“The size effect could not be purely deterministic. Since the material is random, there must exist a component that is statistical, and that component must be of Weibull type,” was another kind of criticism, which I faced at the probabilistic conferences of ASCE and ICOSSAR. Throughout the 1990s, Yunping Xi and Drahoš Novák helped me to clarify this point, both for failures after large crack growth, typical of reinforced concrete, and for failures at crack initiation, typified by the modulus of rupture test. In latter case, this led to an amalgamated energetic-statistical formula for size effect, whose one justification was found in the previous results of Rokugo in Gifu, and later those of Koide in Sendai, Rocco in Madrid, and others.

From the beginning, the computational modeling of distributed fracture in concrete has been one of primary objectives and the main driving forces of progress. I was privileged to be at perhaps the earliest workshop on this subject, organized by Luigi Cedolin in Milano in 1978, under the auspices of Fratelli Pesenti foundation, and two years later at a similar workshop organized by Mang in Vienna.

The early computational models for distributed cracking, quasibrittle fracture and size effect were greatly refined during the 1990s, for example, in the works of Jirásek in Lausanne and Mang in Vienna. Rots in Delft and Červenkas in Prague developed powerful commercial software for distributed

cracking in concrete based on the crack band model (DIANA, SBETA, ATENA). Another practical approach, initiated by Leroy and Ortiz, and refined by Belytschko, Fish, Shing, Jirásek, etc., is to embed discontinuities in the finite elements. A very effective second-gradient model, which transfers the second gradient to a separate differential equation (Helmholtz equation) in order to avoid finite elements of a higher-order continuity, and which can be physically justified as an approximation to the integral-type nonlocal model, was developed by de Borst, Peerlings and other researchers in Delft. The nonlocal and gradient models were recognized as one good way to predict, with sufficient mesh refinement, the crack propagation direction and path. An entirely different approach is the meshless method and partition-of-unity formulations pioneered by Belytschko and his collaborators. The field, however, has become too broad to give proper mention to all of the effective models and contributors.

“Does your nonlocal approach have any physical justification?” was one question thrown at me at Northwestern by Toshio Mura. Initially I had no clear answer, but about a decade ago, the answer materialized in material heterogeneity, the energy release by microcracks, and the interactions of adjacent cracks. The last, however, pointed out the need for a more complex nonlocal model with orientation dependence, which can distinguish mutual crack shielding and amplification. Jirásek and Ožbolt were of great help in calibrating and implementing this model. Thanks to the assistance of Zubelewicz, Tabbara, Jirásek and Cusatis, further physical justification was gained during the 1990s by means simulating the fracture process zone and the size effect in concrete and sea ice with the random particle models and shear-lattice models. However, an effective and universal nonlocal model physically well founded in micromechanics still remains an open problem.

To make sophisticated analyses of distributed cracking and material damage meaningful, realistic constitutive models had to be developed. One fashionable trend, initiated by the late Kachanov and championed by Lemaitre and many others, has been continuum damage mechanics. Most formulations on the abstract side, however, did not include a characteristic length. Some of us in concrete wondered: “Mustn’t this length be inserted in such formulations to make sense for strain softening?” Of course it must, and the models for concrete (e.g., that developed at Northwestern with the help of Pijaudier-Cabot and of Carol, and another formulated in France by Mazars) adhered to this essential requirement.

In an effort to maximize conceptual simplicity and capture physical phenomena such as crack opening and frictional slip in the microstructure, a sequence of progressively refined microplane constitutive models (inspired by a 1938 idea of the great G.I. Taylor) has been developed at Northwestern (first with the help of Oh, then Prat, Hasegawa, Carol, Xiang, Jirásek, Caner and Brocca). What in this approach facilitates capturing the physical phenomena? It is the fact that the microplane constitutive law is formulated not in terms of the stress and strain tensors, but in terms of the stress and strain vectors acting on kinematically constrained planes of all possible orientations in the material, for which the term ‘microplanes’ was coined. “But isn’t the computational work prohibitive?” was a frequent objection. Not any more, thanks to the relentless increase in computer power. At WES in Vicksburg, Mark Adley and Steve Akers have been simulating impact and groundshock with the latest version M4 of the microplane model using an explicit finite element program with several million finite elements. The power of microplane model, originally developed for concrete, is now being exploited for the constitutive modeling of anisotropic rocks, soils, rigid foams, fiber composites, metals and shape-memory alloys (the last with my colleague Cate Brinson).

A difficult challenge, which has so far met with only minor success, has been to convince the engineering community, and especially the code-writing bodies, to take into account the deterministic (energetic) size effect, which is the principal consequence of both the cohesive and crack band models, as well as the nonlocal continuum model. Experimental demonstrations and calibrations were obviously necessary. With this aim, various reduced-scale tests of size effect in concrete have been conducted at Northwestern during the last two decades.

The most impressive tests, however, were made in Japan. At the Japan-US Workshop at the University of Tokyo in 1985, organized by Okamura, we Americans privately wondered: “How come these Japanese companies, Shimizu and Kajima, are so much more enlightened than ours that they would spend all that money on size effect tests of shear failure of prestressed concrete beams and slabs of such enormous sizes (ranging up to 36 m?? and 20 m??, respectively)?” As a result of these tests (reported in 1985 by Iguro, Shiyoa, Nojiri and Akiyama), Japan became the first country to introduce, through Okamura’s persuasion, a formula for size effect into the design code for reinforced concrete. It was also in Japan that the first conference devoted exclusively to size effects in concrete structures was organized—by Mihashi, at Tohoku in 1993. Many innovative practical developments in the modeling of cracking and fracture of concrete structures appeared in Japan, e.g., Horii’s analysis

of shotcrete linings of tunnels, Tanabe's analysis of high temperature exposure, constitutive model implementations by Okamura and Maekawa, etc.

Going by past experience, large concrete dams can be very dangerous to the population downstream, and so they naturally became one problem of heightened concern. During 1990–1994, Wittmann, Saouma and Mazars organized a series of workshops in Lugano, Boulder and Chambéry. The discussions were animated. “Why should we bother about fracture mechanics. We play it safe, treating concrete as a ‘no-tension’ material, for which plasticity-type analysis with no size effect is satisfactory,” objected some dam designers. Yet the need for fracture and size effect analysis was eventually established (most convincingly, in my opinion, through the nonlocal and cohesive finite element studies, particularly those of my former star student Milan Jirásek in Lausanne, and the group of Elices in Madrid). Thanks to the efforts of Victor Saouma in Boulder, a requirement for analyzing cracked dams by fracture mechanics was incorporated into design regulations (and thus the U.S. became probably the first country where this kind of requirement was introduced into mandatory regulations for concrete).

Expanding the deterministic (energetic) size effect characterization from concrete and rock to other quasibrittle materials has been an uphill fight as well. Similar to the earlier situation with concrete, fracture mechanics was in the mid 1980s generally believed to be inapplicable to sea ice, and the only size effect was thought to be statistical. The Navy was interested in navigation in the Arctic and especially in the surfacing of submarines, while the oil companies needed to understand the forces on oil platforms and the load capacity of floating ice. I made efforts to convince the Office of Naval Research (ONR) to start a new initiative on fracture and size effect in sea ice, and then competed with John Dempsey, my friend at Clarkson whom I knew from his previous stay at Northwestern, to get the money. He won the bigger share of it. But we collaborated, and in 1993 I was thrilled to go with John and his large team to the Arctic Ocean to participate, near Cornwallis Island, in his record-breaking size effect experiments in which floating notched ice specimens of dimensions up to $80\text{m}\times 80\text{m}\times 1.8\text{m}$ were fractured under crack opening control. Despite scatter, the measured trend neatly agreed with the simple size effect law. It was a confirmation that sea ice, too, is a quasibrittle material which, on a large scale, obeys fracture mechanics and exhibits a nonstatistical size effect. This changed the field of ice mechanics. The size effect confirmed by John's test results, for example, helped to explain why the measured horizontal forces exerted by moving ice on oil platforms are an order of magnitude smaller than those predicted by elasto-plastic finite element analysis based on the measured laboratory strength of ice. Y.-N. Li and J.-J. H. Kim at Northwestern pursued computer modeling of fracture in vertical penetration of floating sea ice, which confirmed a strong size effect.

The story with fiber-polymer composites and sandwich structures has been similar. Again, fracture mechanics was by the mid 1990s still perceived as inapplicable. It was not used in practice, and was not addressed in textbooks. The size effect was generally either disregarded or regarded as strictly statistical. But the Navy became interested in using these composites and sandwich structures to build large ships, and Yapa Rajapakse at ONR was sceptical of the status quo and very sympathetic to funding a program to get a better grasp of fracture scaling. My colleague and friend Isaac Daniel was a sceptic, too, and so we teamed up. Together with our capable assistants, we again found the size effect in large specimens of fiber composites to be deterministic (energetic), both for failures at crack initiation and those after large crack growth, and in both tension and compression (particularly, in kink-band microbuckling). For sandwich shells, similar conclusions are emerging. Of course, difficult challenges still exist, posed for example by anisotropy, delamination and micromechanics of fiber slip and breakage.

The struggle for improvement in concrete design codes continues today, but the climate has improved. Until recently, the code making committees, consisting mainly of practicing engineers and practice oriented academics, were adamantly opposed to introducing fracture-based size effect formulas, which would imply abandoning the entrenched limit-state design philosophy based on plasticity. Today, though, many members of these committees admit that size effects in concrete structures indeed exist and are important. The size effect is no longer seen as a nuisance, but as a necessity. What is now disputed is which formula to adopt among various proposed alternatives, some purely empirical. Regrettably, there is an enduring tendency to prefer strictly empirical formulas with no theoretical basis over formulas based on some deeper theory (which would have to be learned).

How can we formulate the simple size effect corrections required for design code equations? In my view, they should be derived by asymptotic matching, i.e., smooth ‘interpolation’ between an approximate LEFM-based formula applicable for large sizes and the existing limit-state formula applicable for small sizes. When forced to match the correct asymptotic trends, even if far outside the practi-

cal range, an approximate formula is far more likely to have general applicability, covering practical situations beyond the scope of the test data used for calibration. An empirical formula is unlikely to achieve that, except by sheer luck.

On the probabilistic side, a major problem looms ahead. Based on statistical data, the current safety factors applied to the self-weight of large structures, unlike those applied to the live load, are excessive, and by far. Since the self-weight in large structures contributes a much greater portion of the internal forces than in small ones, the excessive load factor for self-weight (typically 1.4) implies a hidden size effect, which can be as large as 30%.

Doesn't this hidden size effect compensate for the lack of an explicit size effect in the existing limit-state design formulas? In a broad sense, it does, but it does so in an irrational way. To guard against certain kinds of failures (e.g., the ductile bending failure due to yielding of reinforcement), no such compensation is needed, while for others (e.g., the diagonal shear failure or any failure due to crushing of a 'compression strut' in the strut-and-tie model), a much larger compensation may be needed. For high-strength concrete structures or prestressed structures, which are lighter, the size effect compensation hidden in the excessive dead load factor is less than that for normal-strength structures or unprestressed ones, yet the size effect is stronger because of higher brittleness. Simply, burying the size effect in the dead load factor is not a smart idea.

At the 'fib' congress in Prague, a noted engineer asked in a doubtful tone: "Has the size effect ever caused any structural failures?" It surely has, but it has not been the exclusive cause. The size effect (ranging from 25% to 55%) has likely been a significant contributing factor in many famous catastrophes, for example, the failures of Malpasset dam in the Alps (1959), St. Francis dam in California (1928), Schoharie Creek Bridge on the New York Thruway (1987), Sleipner oil platform in Norway (1991), Han-Shin viaduct in the Kobe earthquake (1995), Cypress viaduct in Oakland in the Loma Prieta earthquake (1989), and the bridge columns in Los Angeles in the Northridge earthquake (1984). So why, with the exception of the Schoharie Creek bridge, has the size effect not been recognized by the investigating committee of experts? The reason must be that the safety factors in civil engineering, in contrast to aeronautical engineering, are very large. Normally several kinds of errors must conspire to bring down a bridge or building, while only one error (perhaps such as neglecting to take into account the size effect in a large aircraft tail made entirely of fiber composites) suffices to doom an airliner.

Let me end by returning to the statistics of self-weight errors. Should we accept the recent proposals to reduce the dead load factor to a much lower value, as justified by statistics? Absolutely not. Such a reduction would be dangerous unless the size effect would at the same time be incorporated into the code provisions. Vice versa, it must be admitted that it would make little sense to introduce the size effect into the code without reducing the dead load factor for the self-weight to a realistic value.

Likewise, the purpose of a detailed finite element analysis of fracture and size effect in a large concrete structure is defeated if one applies the excessive dead load factor prescribed by current codes. With that kind of irrational safety factor, a simple back-of-the-envelope calculation might be just as good.

A computational enthusiast suggested to me: "Why don't you ignore the safety factors in the code and turn to the stochastic finite elements?" That is a tempting proposition, but it is not feasible at present. The existing stochastic finite element formulations, in their current state, cannot predict a load of extremely small failure probability, such as 10^{-7} , because the tail structure of the probability distribution is not correct (to wit, if these formulations were correct, then, for the case of a structure failing at fracture initiation, exemplified by the modulus of rupture test, they would have to converge to the Weibull-type size effect, but they do not). What's the problem? The distribution tails are essentially exponential, while to be correct they must be of Weibull type, as dictated by extreme value statistic. It would be necessary to devise a stochastic finite element method in which the probability of the first eigenvalue of the tangential stiffness matrix being nonpositive becomes Weibull distributed as failure is approached. This is still an open problem. At present, stochastic finite elements can realistically predict loads with failure probability about 1 to 10%, but hardly much less. Obviously, synergistic efforts by fracture experts, statistical reliability experts, and computational mechanics experts are needed.

The struggle is far from over, but further progress will make it worthwhile.