Prolegomena to the Study of Friction Stir Welding

A. C. Nunes, Jr.
Materials and Processes Laboratory, Marshall Space Flight Center
Huntsville, AL, USA

Abstract

The literature contains many approaches toward modeling of the friction stir welding (FSW) process with varying treatments of the weld metal properties. It is worthwhile to consider certain fundamental features of the process before attempting to interpret FSW phenomena: Because of the unique character of metal deformation (as opposed to, say, viscous deformation) a velocity “discontinuity” or shear surface occurs in FSW and determines much of the character of the welding mechanism. A shear surface may not always produce a sound bond. Balancing mechanical power input against conduction and convection heat losses yields a relation, a “temperature index”, between spindle speed and travel speed to maintain constant weld temperature. But many process features are only weakly dependent upon temperature. Thus, unlike modeling of metal forming processes, it may be that modeling the FSW process independently of the material conditions has some merit.

Introduction

The title word “prolegomena”, from the Greek meaning “things said before”, or “prefatory remarks” is often used in a philosophical context as in Immanuel Kant’s Prolegomena to Any Future Metaphysics or Jane Harrison’s Prolegomena to the Study of Greek Religion, etc.. It was chosen to evoke a philosophical mood. The purpose of this paper is to share with the friction stir welding community a set of concepts for the interpretation of weld structures and observations that have been found helpful and to subject them to their scrutiny.

The Difference between Plasticity and Viscosity

The dependence of flow stress on shear rate is strong in viscous media and weak in plastic media. This is clear from kinetic models of Fig. 1 and 2 and empirical data. Hence velocity “discontinuities” or “adiabatic” shear zones are anticipated in plastic media like metals, for example in metal cutting and friction stir welding, but not in viscous media.

Fig. 1 illustrates a hypothetical deformation process in which shear takes place through atomic-level rearrangements or jumps. By symmetry it is presumed that there are as many sites ready to contribute a positive as a negative shear increment. Tentative jumps are made at frequency $f$, the natural oscillation frequency or the Debye frequency, but only succeed if the kinetic energy of the attempting atom is sufficient to clear the highest intermediate potential energy configuration in its way, i.e. the activation energy $E$. 
for the jump. The probability that a jumping atom has this amount of kinetic energy is 
\[ P = e^{-\frac{E}{k_B T}} \], where T is the absolute temperature and \( k_B \) is Boltzmann’s constant.

If a shear stress acts on the body, it is transmitted to the deformation locality and unbalances the back and forth jump rates to favor a net shear deformation rate \( \dot{\gamma} \). The shear stress reduces the activation energy for deformations in the forward direction and raises the activation energy for deformation in the reverse direction. Using the geometry of Fig. 1 an approximate relation between strain rate and stress can be obtained.

\[
\dot{\gamma} \approx b \left( \frac{\Delta x \Delta y}{L^2} \right) \frac{1}{L} \left[ \frac{1}{2} e^{-\frac{b \Delta x \Delta y}{k_B T}} - \frac{1}{2} e^{\frac{b \Delta x \Delta y}{k_B T}} \right]
\]

\[ = b \left( \frac{\Delta x \Delta y}{L^2} \right) \frac{1}{L} \left[ \frac{1}{2} e^{-\frac{E}{k_B T}} \right] \left[ 1 + \frac{b \Delta x \Delta y}{k_B T} \sinh \left( \frac{b \Delta x \Delta y}{k_B T} \right) \right] \]  

(1)
The competing forward and reverse deformations results in a net dependence of shear rate upon the hyperbolic sine of the stress. If $\tau b \Delta x \Delta y / k_BT << 1$, which will at least be true at the incipient application of stress, $\sinh \left( \frac{b \Delta x \Delta y}{k_BT} \right) \rightarrow \frac{b \Delta x \Delta y}{k_BT}$, and the stress-strain-rate equation (1) becomes that of a viscous medium,

$$\tau \sim \mu \dot{\gamma} \quad (2)$$

where the viscosity $\mu$ is given by

$$\mu \sim \frac{2L^3k_BT}{f_n(b\Delta x \Delta y)^2}e^{E/k_BT} \quad (3)$$
It is possible to insert some numbers into equation (3) to see how close this crude approximation may come to a realistic value. For molten aluminum a viscosity on the order of a few centipoise would be anticipated.

Suppose \( b \sim \Delta x \sim \Delta y \sim 3 \text{Å}, \ L \sim 1,000 \text{Å}, \ n \sim 10^6, \ f \sim 8 \times 10^{12} \text{sec}^{-1} \)

\[
 k_B = \frac{1.38 \times 10^{-23}}{\text{W} \cdot \text{sec} \cdot \text{K}}, \quad k = 140 \frac{\text{W}}{\text{m} \cdot \text{K}} \cdot 10^{-10} \frac{\text{m}}{\text{Å}} = 1.4 \times 10^{-8} \frac{\text{W}}{\text{Å} \cdot \text{o K}}, \quad \text{and}
\]

\[
 E \sim 10^{-20} \text{W} \cdot \text{sec}. \quad \text{Then}
\]

\[
 \mu \sim 4.73 \times 10^{-3} \frac{T^{25}}{T} \text{ centipoise} \quad (4)
\]

At \( T = 660^\circ \text{C} \), \( \mu \sim 3.43 \) centipoise.

Fig. 2 illustrates a hypothetical deformation process in which shear takes place through breakthroughs of dislocations past pinning point barriers. Attempts are made at
frequency \( f \), the natural oscillation frequency, to break through a pinning point. Attempts only succeed if the kinetic energy of the attempting atomic-scale structure at the pinning point is sufficient to clear the highest intermediate potential energy configuration in its way, i.e. the activation energy \( E \) for the jump. The probability that a structure has this amount of kinetic energy is 
\[ e^{-\frac{E}{k_B T}} \]
where \( T \) is the absolute temperature and \( k_B \) is Boltzmann’s constant.

If a shear stress acts on the body, it is transmitted to the pinning point locality and the shear stress adds to that of any local shear stress fluctuation due to thermal oscillations. The shear stress reduces the activation energy for deformation in the forward direction. Since the dislocation has to be pressed against a pinning obstacle for there to be a potential jump site, the number of reverse jump sites is taken to be zero. (This feature also causes the model to break down for very low stresses and strain rates.) For the deformation mechanism depicted in Fig. 2 an approximate relation between strain rate and stress can be obtained.

\[ \dot{\gamma} \sim b \left( \frac{\Delta a}{L^2} \right) \left( \frac{1}{L} \right) \frac{E - \left( \frac{b \Delta x \Delta s}{2} \right)}{k_B T} e^{-\frac{E - \left( \frac{b \Delta x \Delta s}{2} \right)}{k_B T}} \]

(5)

According to equation (5) the mechanism of Fig. 2 exhibits a weak stress dependence upon the logarithm strain rate rather than the strong linear dependence of equation (2) for the mechanism of that of Fig. 1.

\[ \tau \sim \frac{E}{b \Delta x \Delta s} + \frac{k_B T}{b \Delta x \Delta s} \ln \left[ b \left( \frac{\Delta a}{L^2} \right) \left( \frac{1}{L} \right) n f e^{-\frac{E - \left( \frac{b \Delta x \Delta s}{2} \right)}{k_B T}} \right] \]

(6)

Suppose \( b \sim \Delta x \sim 3\AA, \Delta s \sim 200b, \Delta a \sim \frac{\pi}{4} \Delta s^2 \sim 10,000\pi b^2, L \sim 1,000\AA, n \sim 0.1, f \sim 8 \times 10^{12} \text{sec}^{-1}, k_B = 1.38 \times 10^{-23} \text{J} \cdot \text{K}^{-1}, k \sim 140 \frac{\text{W} \cdot \text{m} \cdot \text{s}}{\text{K}^2}, E \sim 0.2G \text{b}^3, \text{and } G \sim 4 \times 10^6 \text{psi},

\[ \tau \sim 16,000 - 16.5T \text{ psi} \]

(7)

At \( T \sim 450^\circ \text{C} \) and \( \dot{\gamma} \sim 10^4 \text{sec}^{-1}, T \sim 4,070 \text{ psi}. \)
Singer and Cottrell [1] have measured a stress-temperature dependence similar to that of equation (7) for Al-Si alloys close to the solidus temperature, the temperature range most pertinent to friction stir welding. The stress-temperature dependence over a wide temperature range appears rather different because it represents a composite of different deformation mechanisms.) This could be the consequence of a solid solution hardening mechanism. A similar dependence is also found [2] in pure aluminum at temperatures in the vicinity of absolute zero, although the dislocation obstacles and the deformation mechanism parameters are different.

Since the 1930’s [2] the plastic deformation of metals has been understood in terms of dislocation theory. While the above deformation model is too crude to do more than illustrate some trends, it implies that plastic deformation of metals is quite distinct from viscous deformation. A glance at some old data [3] for mild steel confirms this. The yield point of a mild steel at room temperature varies from 45 ksi at a strain rate of 1.0 per second rises only to 80 to 86 ksi at $10^3$ per second. The strain rate rises by a factor of 1000, but the yield stress by a factor of less than 2. No wonder that it is customary to neglect the strain-rate dependence of the flow stress in plasticity theory [4] and to treat the flow stress as a constant where the temperature remains fixed. But the insensitivity of the flow stress of metals to strain rate has an important consequence for the friction stir welding process.

Let us consider a block of material of dimensions $\Delta x$ by $\Delta y$ by $\Delta z$ subjected to a shear stress $\tau$ on the $\Delta x$ by $\Delta y$ plane as shown in Fig. 3.

Equilibrium requires
\[ \rho \frac{\partial v}{\partial t} = \frac{\partial \tau}{\partial z} \quad (7) \]

Conservation of energy requires
\[ \rho C \frac{\partial T}{\partial t} = k \left( \frac{\partial^2 T}{\partial z^2} \right) + \tau \frac{\partial v}{\partial z} \quad (8) \]

Recalling that \( \dot{\gamma} = \frac{\partial v}{\partial z} \), the constitutive equation depends on the material. For the viscous material, taking \( A = \frac{fn(bA\Delta y)^2}{2L^3k_B} \) as a constant
\[ \frac{\partial v}{\partial z} = A e^{-\frac{E}{k_B T}} \frac{\tau}{T} \quad (9) \]

For the plastic, metallic material, taking \( B \) as a constant and labeling the term from equation (5) \( \frac{bA\Delta x\Delta s}{2} = v_a \), the “activation volume”
\[ \frac{\partial v}{\partial z} = Be^{-\frac{E}{k_B T}} e^{\frac{v_a}{k_B T}} \quad (10) \]

We note in passing that if the temperature and shear stress remain constant over \( z \), so does the rate of shear according to the constitutive equations. In general, however, the temperature will vary over \( z \) and so will \( v \). For quasi-steady conditions, where the time derivatives vanish,
\[ \frac{\partial \tau}{\partial z} = 0 \quad (11) \]

and the shear stress remains constant over \( z \). We note that \( \dot{\gamma} = -\frac{31.5}{\tau} \left( \frac{d^2T}{dz^2} \right) \) sec\(^{-1} \) and \( v = v_o - \frac{31.5}{\tau} \left( \frac{dT}{dx} \right) \) in/sec.

For the viscous material
\[
\frac{d^2T}{dz^2} = -\left( \frac{A\tau^2}{k} \right) e^{\frac{-E}{k_B T}} \tag{12}
\]

or, inserting the above numbers,

\[
\frac{d^2T}{dz^2} = -\frac{4632\tau^2}{T} e^{-T/725} \tag{13}
\]

and at \(\tau = 4\) psi and initial \((z = 0)\) temperature of \(450^\circ C\), a numerical solution for the velocity distribution \((v_o = 0)\) yields Fig. 4.

\[\text{Figure 4 Numerical solution for viscous velocity distribution at shear stress of 4 psi with initial temperature 450^\circ C}\]

Fig. 4 presents an example of homogeneous shear.

For the plastic material with \(B = b \left( \frac{\Delta a}{L^2} \right) \left( \frac{1}{L} \right) \eta\)

\[
\frac{d^2T}{dz^2} = -\left( \frac{B\tau}{k} \right) e^{\frac{E-\tau v}{k_B T}} \tag{14}
\]

or, inserting the numbers from above

\[
\frac{d^2T}{dz^2} = -2.155 \times 10^7 \tau e^{\frac{10,790 - 0.6744 \tau}{T}} \tag{15}
\]
and at $\tau = 4,000$ psi and initial ($z = 0$) temperature of $450^\circ$C, a numerical solution for the velocity distribution yields Fig. 5.

![Figure 5 Numerical solution for plastic velocity distribution at shear stress of 4,000 psi with initial temperature 450°C](image)

Fig. 5 presents an example of inhomogeneous shear. In metal processing inhomogeneous deformation is often a concern to be avoided [6]. The shearing takes place within a narrow (0.02 inch) band. The average rate of shear within this “adiabatic shear band” is on the order of 4,500 sec$^{-1}$.

**Friction Stir Welding Compared to Metal Cutting**

The above is a very crude demonstration of how atomic-level deformation mechanisms give rise to constitutive relations of materials, which, in turn, give rise to macroscopic deformation modes. Adiabatic shear bands like that computed for Fig. 5, but appreciably thinner, observed around the friction stir pin-tool play a critical role in the success of the welding process. They separate a rotating plug of metal stuck to the tool from the (almost) undeforming metal surrounding the tool (see Fig. 6).

Another place where adiabatic shear bands present themselves is in metal cutting [7] and an analogy between the metal cutting process and the friction stir welding process is illustrated in Fig. 6. The friction stir process resembles a pair of metal cutting operations back-to-back with the metal in the rotating plug corresponding to the metal cutting chip. The adiabatic shear surface metal typically appears periodically in metal cutting chips. In friction stir welding the process constraints result in adiabatic shear surface, but even here there is an oscillation. The friction stir oscillations give rise to the so-called tool marks on the weld surface following the tool and to textural bands [8] in the interior of the weld, for example the “onion rings” in the weld “nugget”. The cause of these oscillations has not yet been established. They could be the result of a self-excited process characteristic of the flow or the result of a forced oscillation due to tool
asymmetry. The usual synchronization of the oscillations with the tool rotation period suggests a forced oscillation, but it is not certain that this is always the case.

As in metal cutting, the adiabatic shear band in friction stir welding is surrounded by a broader deformation field due to other, much slower auxiliary deformation processes. For many purposes, for example interpreting weld microstructures, auxiliary deformation processes can be neglected and the kinematics of the welding process can be modeled independent of the material conditions. If the slip is primarily on an approximately fixed shear surface, then an input-output power balance can be constructed easily and can determine weld temperature, constant temperature isotherms for parameters, and indices for temperature, power input, or energy input as is demonstrated below.

Figure 6 Analogy between friction stir welding and metal cutting
The Nature of the Welding Process

Putting clean surfaces into intimate contact makes solid-state welds. This requires a pressure high enough to push down surface asperities and a clean surface free of ordinary contamination by dirt and oxides. The high shear induced at the friction stir weld shear zone stretches out the seam to be welded so far as to render insignificant the original contaminated surface as shown in Fig. 7, and, given adequate pressure, to ensure a sound bond. If the weld metal passes through the shear zone in time increment $\Delta t$, the initial seam length $L_0 = V \cos \theta \Delta t$ is expanded by the rotation of the plug to $L \approx R \omega \Delta t$, hence the effective elongation of the seam is

$$\frac{L - L_0}{L_0} \approx \frac{R \omega}{V \cos \theta} - 1 \approx \frac{R \omega}{V \cos \theta}$$

(16)

In situations where the seam does not pass through the shear surface, for example seams far to the retreating side of the weld or plug welds, the effective seam surface is not stretched out and the contamination level is not significantly reduced for bonding, and the bond quality is problematic. See Fig. 8.

![Figure 7 Strain at the shear surface and the friction stir welding mechanism](image)
Figure 8 Conditions under which weld seams are not subjected to high strain rate at a shear surface

Parameter Temperature Window

Excessively cold welds cause high tool stresses. Excessively hot welds are accompanied by problems with metal flow out from under the tool shoulder (e.g. flash, trenching). A heat balance between mechanical power input and thermal conductive and convective heat losses determines weld temperature. See Fig. 9.

A simplified quantitative model demonstrates trends. The model also yields a relation between RPM and weld speed for constant weld temperature (a “temperature index”).

\[
\frac{2\pi R^2 W \omega}{\text{Mechanical Power Input}} = \frac{2\pi kW}{\ln\frac{R}{R_0}}(T - T_0) + 2RwV\rho C(T - T_0)
\]

(18)
The heat balance of equation (18) can be rewritten as in equation (19)

\[ V \approx -\frac{\pi k}{\rho C R \ln \frac{R_o}{R}} + \frac{\pi R \tau (T)}{\rho C (T - T_o)} \omega \]  

(19)

Equation (19) relates the two principal friction stir weld parameters, weld speed \( V \) and RPM \( \frac{60 \omega}{2\pi} \) if \( \omega \) is in units of radians/second) for a rotating cylinder. For metals the torque depends only weakly on shearing rate, hence we approximate the shear stress \( \tau \) as a function of temperature only. Within this approximation for constant weld temperature the coefficient of \( \omega \) is constant given the same weld metal (\( \rho \) is the weld metal density and \( C \) is its specific heat.) and geometry (\( R \) is the cylindrical shear surface radius.). Likewise the preceding term is a constant. One can anticipate that the isotherm parameter relations are approximately a family of straight lines fanning out from a point as shown in Fig. 10.
Temperature Index

If the form of the shear stress relationship to temperature is $\tau \approx a - bT$, where $a$ and $b$ are constants, then it is possible to construct an approximate temperature index from equation (19):

$$\frac{T - T_o}{T_o} \approx \left( \frac{a}{bT_o} - 1 \right) \frac{\pi b}{\rho C} \left( \frac{R \omega}{V} \right)$$

$$1 + \left( 1 + \frac{k}{bR^2 \omega \ln \frac{R_o}{R}} \right) \frac{\pi b}{\rho C} \left( \frac{R \omega}{V} \right) \frac{R_o}{V}$$

(20)
At very slow speeds ($\frac{R\omega}{V} \gg 1$) using the values from above computations with $R = 0.125$ inches a temperature index depending upon RPM alone is obtained:

$$\frac{T - T_o}{T_o} \approx \frac{2.25}{1 + \frac{40.8}{\omega}} = \frac{2.25}{1 + \frac{390}{\text{RPM}}}$$

(21)

The temperature index is 0.763 ($T \sim 252^\circ C$, $\tau \sim 7,340$ psi) for 200 RPM, 1.14 ($T \sim 364^\circ C$, $\tau \sim 5,490$ psi) for 400 RPM, up to 2.25 ($T \sim 696^\circ C$, $\tau \sim 12$ psi) as the RPM approaches infinity. When weld speed cannot be neglected

$$\frac{T - T_o}{T_o} \approx \frac{0.257 \left( \frac{\text{RPM}}{V} \right)}{1 + \left( 0.1144 \right) \left[ 1 + \frac{390}{\text{RPM}} \left( \frac{\text{RPM}}{V} \right) \right]}$$

(22)

For a weld speed $V = 10$ inches per minute, at 200 RPM the temperature index is 0.663 ($T \sim 223^\circ C$, $\tau \sim 7,820$ psi); at 400 RPM, 1.02 ($T \sim 329^\circ C$, $\tau \sim 6,070$ psi); and at infinitely high RPM, 2.25 ($T \sim 696^\circ C$, $\tau \sim 12$ psi).

To change parameters but maintain constant weld temperature, the temperature index needs to be maintained constant.

$$0.257 \left( \frac{\text{RPM}_1}{V_1} \right) \left[ 1 + \left( 0.1144 \right) \left( \frac{390}{\text{RPM}_1} \left( \frac{\text{RPM}_1}{V_1} \right) \right) \right] = 0.257 \left( \frac{\text{RPM}_2}{V_2} \right) \left[ 1 + \left( 0.1144 \right) \left( \frac{390}{\text{RPM}_2} \left( \frac{\text{RPM}_2}{V_2} \right) \right) \right]$$

(23)

**Other Indices**

Although temperature is, perhaps, the most important criterion for weld equivalence, it is possible to construct a heat index, i.e. a measure of the power entering the weld, or an energy index, a measure of the energy received per unit length of weld.

From equations (18) and constitutive relation $\tau = a - bT$ it is possible to estimate the power input $P$ and to use it, or a function thereof, as a heat index.
\[
P \approx 2\pi R^2 \eta \left(\frac{a}{b} - T_\circ \right) \left(\frac{2\pi k_w}{\ln \frac{R_o}{R}} + 2R w V \rho C\right) \left(1 + \frac{2\pi k_w}{\ln \frac{R_o}{R}} + 2R w V \rho C\right)
\]

The energy per unit length of weld, the basis for an energy index will then be

\[
E = \frac{P}{V}
\]  

(25)

**Conclusion**

These very rough computations demonstrate in a simple way how:

1. The constitutive equations of materials depend upon their underlying structure and that the structure and constitutive equations of metals are distinct from viscous media.

2. The flow characteristics of media depend upon their constitutive equations. The shearing region in metals may, under appropriate conditions, contract to an adiabatic shear zone such as is seen in metal cutting and in friction stir welds.

3. The adiabatic shear zone plays a critical role in the mechanism by which friction stir welds are made. The existence and geometrical configuration of the adiabatic shear zone depends only weakly on material conditions, hence a kinematic model of the weld flow field can be constructed more or less independent of material conditions.

4. There are situations where the weld seam does not encounter the adiabatic shear zone and full weld strength is not achieved.

5. A simplified heat balance using the adiabatic shear zone concept determines the approximate form of weld isotherms on a plot of weld speed vs. RPM. A too-cold and a too-hot isotherm will bound a window of acceptable parameters, however, the window may be further constricted by various defect formation mechanisms.

6. A temperature index (or a heat index or an energy index) can be constructed on the basis of this heat balance, such that keeping the index constant keeps the temperature (or heat input or energy input) constant when the weld RPM and speed parameters are changed.
Of course the computations presented can be elaborated and made more precise. But even in the rough form presented herein they provide a basis for interpreting many empirical observations.

References


